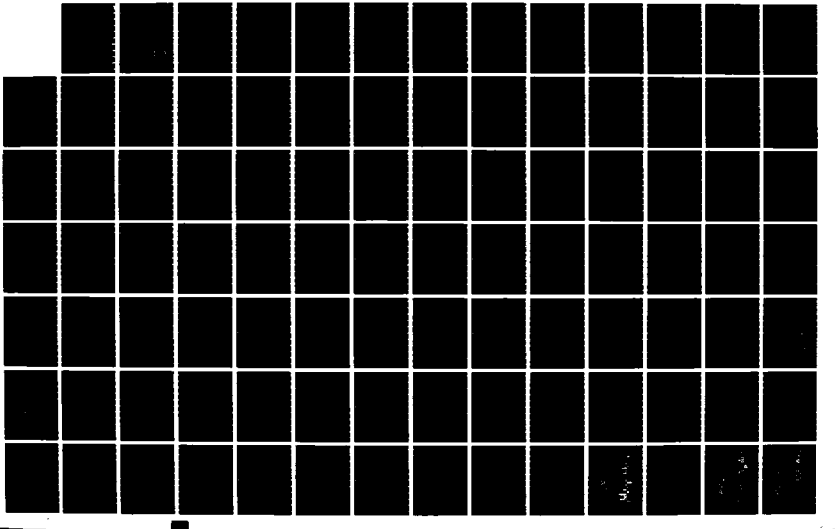
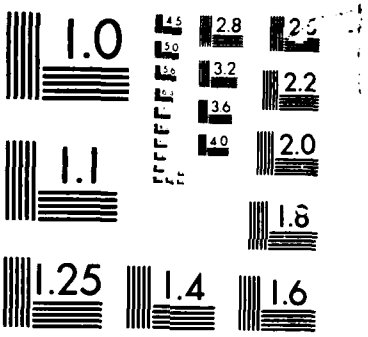
AD-A168 897 RECONSTRUCTION OF UNDERSAMPLED PERIODIC SIGNALS(U) 1/2
MASSACHUSETTS INST OF TECH CAMBRIDGE RESEARCH LAB OF
ELECTRONICS A J SILVA JAN 86 TR-514 N00014-81-K-0742
UNCLASSIFIED F/G 9/3 NL




1.0
1.1
1.25
1.4
1.6
2.8
3.2
3.6
4.0
2.2
2.0
1.8

Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
Research Laboratory of Electronics
Room 36-615
Cambridge, MA 02139

Reconstruction of Undersampled Periodic Signals

Anthony J. Silva

Technical Report No. 514

January 1986

This work has been supported in part by the Advanced Research Projects Agency monitored by ONR under Contract No. N00014-81-K-0742 and in part by the National Science Foundation under Grant ECS-8407285.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Research Laboratory of Electronics Massachusetts Institute of Technology 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Office of Naval Research Mathematical and Information Scien. Div.
6c. ADDRESS (City, State and ZIP Code) 77 Massachusetts Avenue Cambridge, MA 02139	7b. ADDRESS (City, State and ZIP Code) 800 North Quincy Street Arlington, Virginia 22217
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Advanced Research Projects Agency 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-81-K-0742
8c. ADDRESS (City, State and ZIP Code) 1400 Wilson Boulevard Arlington, Virginia 22217	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. NR 049-506 / WORK UNIT NO.
11. TITLE (Include Security Classification) Reconstruction of Undersampled Periodic Signals	
12. PERSONAL AUTHOR(S) Anthony J. Silva	
13a. TYPE OF REPORT Technical 13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day) January 1986 15. PAGE COUNT 118
16. SUPPLEMENTARY NOTATION	
COSATI CODES: FIELD / GROUP / SUB. GR. 17. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Under certain conditions, a periodic signal of unknown fundamental frequency can still be recovered when sampled below the Nyquist rate, or twice the highest frequency present in the waveform. A new sampling criterion has been proposed which enumerates such conditions. It has been shown that in theory, if the signal and sampling frequencies are not integrally related, and the signal is band-limited (to a range the extent of which is known but otherwise unrestricted), then the signal waveshape can always be recovered. If the fundamental frequency is known to lie within a range not spanning any multiple of half the sampling rate, then the temporal scaling for the reconstructed waveform can be determined uniquely, as well. Procedures have also been proposed for reducing time-scale ambiguity when the latter condition is not met.

(cont.)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ OTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Kyra M. Hall RLE Contract Reports	22b. TELEPHONE NUMBER (Include Area Code) (617) 253-2569 22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

19. Abstract continued

A previously presented time domain algorithm for reconstructing aliased periodic signals has been implemented and modified. A new algorithm, operating in the frequency domain, has been proposed and implemented. In the new algorithm, the signal fundamental frequency is first estimated from the discrete Fourier transform of the aliased data through an iterative procedure. This estimate is then used to sort the aliased harmonics. The inverse discrete Fourier transform of the resulting spectrum provides the reconstructed waveform, corresponding to one period of the original signal. Empirical analysis has indicated that the proposed algorithm is comparable to the time domain algorithm in terms of reconstruction quality, robustness, and efficiency.

Reconstruction of Undersampled Periodic Signals

by

Anthony J. Silva

Submitted to the

Department of Electrical Engineering and Computer Science

on January 31, 1986 in partial fulfillment of the requirements

for the Degree of Master of Science

Abstract

Under certain conditions, a periodic signal of unknown fundamental frequency can still be recovered when sampled below the Nyquist rate, or twice the highest frequency present in the waveform. A new sampling criterion has been proposed which enumerates such conditions. It has been shown that in theory, if the signal and sampling frequencies are *not integrally related*, and the signal is band-limited (to a range the extent of which is known but otherwise unrestricted), then the signal waveshape can always be recovered. If the fundamental frequency is known to lie within a range not spanning any multiple of half the sampling rate, then the temporal scaling for the reconstructed waveform can be determined uniquely, as well. Procedures have also been proposed for reducing time-scale ambiguity when the latter condition is not met.

A previously presented time domain algorithm for reconstructing aliased periodic signals has been implemented and modified. A new algorithm, operating in the frequency domain, has been proposed and implemented. In the new algorithm, the signal fundamental frequency is first estimated from the discrete Fourier transform of the aliased data through an iterative procedure. This estimate is then used to sort the aliased harmonics. The inverse discrete Fourier transform of the resulting spectrum provides the reconstructed waveform, corresponding to one period of the original signal. Empirical analysis has indicated that the proposed algorithm is comparable to the time domain algorithm in terms of reconstruction quality, robustness, and efficiency.

Thesis Supervisor: Alan V. Oppenheim

Title: Professor of Electrical Engineering

Dedication

To the memory of my father,

Anthony D. Silva.

Thanks for providing me with the

opportunities you never had.

Acknowledgments

I first would like to thank my thesis advisor, Prof. Alan V. Oppenheim, for the encouragement and intellectual stimulation he has provided. No single person has had a greater effect on my professional development. With the exception of my wife-to-be and members of my immediate family, the same can be said of him concerning my personal growth, as well. I would also like to thank my undergraduate thesis advisor, Prof. Campbell L. Searle, for his guidance and encouragement during the earlier part of my graduate career. The impact of the sound advice he gave me at several critical times cannot be overestimated.

I am greatly indebted to Mr. Charles M. Rader at the M.I.T. Lincoln Laboratory for his suggestion of the research topic, and for reviewing the thesis manuscript. While credit for the development of the time domain de-aliasing algorithm described in this document belongs to Charles, any errors or omissions are exclusively my own.

Many of the burdens commonly associated with graduate study have been virtually eliminated by the generous support of my employer, RCA/Automated Systems Division, in Burlington, Massachusetts. I would like to thank Messrs. Eugene M. Stockton, Andrew T. Hospodor, and David M. Priestley for providing me with the opportunity to participate in the RCA Graduate Studies Program which furnished this support. Mr. George W.K. Mukai deserves thanks for his suggestions and instructions for producing the high quality figures in this report within a reasonable amount of time.

Several members of the M.I.T. Digital Signal Processing group have been instrumental in converting my dread of computers to fanaticism, and for resolving differences of opinion between my new love and me when necessary.

My fiancée, Almerinda Gomes, and my mother, Mary Silva, deserve special mention

for their love, support, and toleration of my daily mood swings throughout the years I have spent at M.I.T.

Contents

1 Introduction 1

1.1 Nature of the Problem 1

1.2 Background 2

1.3 Scope, Contribution, and Organization of the Thesis 4

2 Development of a Sampling Criterion for Periodic Signals 5

2.1 The Nyquist Sampling Criterion 5

2.2 The Pseudo-Nyquist Sampling Criterion 7

2.3 Reducing Signal Fundamental Frequency Ambiguity 13

3 Rader Time Domain Sample Sorting Algorithm 17

3.1 General Approach 18

3.2 Detailed Description of the Algorithm 24

3.3 A Modification of the Algorithm 37

3.4 Examples 48

4 SPEC-PEAKS — A Frequency Domain Alternative to the Rader Algorithm 59

4.1 General Approach and Detailed Description of the Algorithm 60

4.2 An Enhancement of the Algorithm 68

4.3 Examples 73

5 Analysis and Conclusions 79

5.1 Reconstruction Quality and Algorithm Robustness 80

5.2 Algorithm Efficiency 95

6 Suggestions for Future Research **102**

List of Figures

2.1 Procedure for recovering an aliased signal. 11

3.1 Formation of a composite period. 19
3.2 Composite periods from various quantities of samples. 23
3.3 Ambiguity of variation. 25
3.4 Hypothetical variation function. 26
3.5 Generation of a Farey series. 28
3.6 Program RADER. 38
3.7 Subroutine PS-NYQ-CRIT. 40
3.8 Subroutine INIT-FF-SEQ. 41
3.9 Subroutine VARIATION. 43
3.10 Subroutine MULT-INVERSE. 44
3.11 Subroutine RECONSTRUCT. 45
3.12 Program FAST-SCAN. 49
3.13 Subroutine RADER-SRCH. 50
3.14 Subroutine RAISE-INIT. 52
3.15 Aliased sinewave recovered using Rader algorithm. 53
3.16 Aliased synthetic signal recovered using Rader algorithm. 55
3.17 Aliased line interference signal recovered using Rader algorithm. 56
3.18 Convergence of variation function. 57

4.1 Computation of partial energy. 67
4.2 Program SPEC-PEAKS. 69
4.3 Subroutine MAX-HARM-E. 70

4.4 Subroutine SORT-HARM. 71
4.5 Adjustment of estimated harmonic locations. 72
4.6 Subroutine ADJ-HARM. 74
4.7 Aliased synthetic signal recovered using SPEC-PEAKS. 76
4.8 Aliased line interference signal recovered using SPEC-PEAKS. 77

5.1 Poor reconstructions when pseudo-Nyquist criterion is not met. 81
5.2 Reconstructions of discontinuous waveform. 83
5.3 Reconstructions when relative harmonic amplitudes change. 84
5.4 Poor reconstructions when fundamental frequency changes. 86
5.5 Reconstructions of noisy waveform. 87
5.6 Reconstructions of two superimposed waveforms. 90
5.7 Periodicity of variation function. 97
5.8 Rader algorithm reconstructions using different search ranges. 98

List of Tables

5.1 Estimation of ϕ_w from two superimposed waveforms. 93

5.2 Rader algorithm recovery time vs. number of input samples. 95

5.3 Rader algorithm search time vs. search range. 100

5.4 SPEC-PEAKS recovery time vs. number of significant harmonics. . . . 101

Chapter 1

Introduction

1.1 Nature of the Problem

In many instances, knowledge of some special property of an analog signal can be exploited to reduce the sampling rate or the number of samples necessary to retain all the information in the signal. Nyquist sampling of bandlimited signals certainly represents one example. As another example, it might be known that the waveform under observation corresponds to one of only a few candidates, and therefore relatively few samples are needed to identify it uniquely. In an extreme case, the signal is known completely beforehand to within a scale factor, in which case only one sample is needed.

In this thesis, we shall first propose a set of sufficient conditions under which a *periodic* signal can still be recovered after uniform sampling below the Nyquist rate, or twice the frequency of the highest harmonic present in the waveform. Next, we will discuss, implement, and modify a time domain algorithm developed by Rader [1] for determining the period of such waveforms and reconstructing them from the samples. For brevity, hereafter we will refer to the combination of these two steps as *de-aliasing*, under the assumption that only periodic signals will be treated. A new frequency domain de-aliasing algorithm will then be developed, and it will be compared with the Rader algorithm.

The work summarized in this thesis should have practical significance since periodic signals abound in both natural and synthetic environments, and it is not always possible

to sample them above the Nyquist rate. While undersampling is typically due to the physical limitations of the available sampler hardware, there are others reasons, as well. It might be desirable to use hardware configured for a low frequency application to sample infrequent or unanticipated high frequency or harmonically rich periodic signals, as may be the case in a satellite in space. Undersampling might be desired for purely economic reasons, since high-speed sampling systems are relatively expensive. The savings would be even greater if it was necessary to sample several periodic signals (whose frequencies need not be related) concurrently, or at least nearly so. A single commutating sampler could be used if the effects of undersampling could be removed at a later time. Applications in bandwidth compression of periodic signals are also possible.

The algorithms to be presented have the benefit of being insensitive[1] to the bandwidth of the original signal, i.e., to the extent of the frequency range containing all signal harmonics. This is a significant advantage over methods such as those comprising decomposition of wide-band signals into several narrow-band components, sampling (at a low rate), and subsequent recombination of the samples to yield a sequence which is not aliased. Multiple samplers are required for such methods, and their number is proportional to the total bandwidth.

It should be emphasized that the goal of this research is to yield solutions in situations where undersampling is unavoidable, or desirable for reasons similar to those mentioned above. It is the minimum sampling rate and not the minimum number of samples necessary that we wish to reduce.

1.2 Background

Signal reconstruction from corrupted data has been and remains a popular topic in discrete-time signal processing. Techniques for removing or reducing noise, reverberation, and other such degradations have been implemented successfully in many instances. However, relatively little work has been published on removing the distor-

[1] At least in theory, and for the most part, in practice as well.

tion introduced by undersampling.

Marks [2] has provided a closed-form method for recovering any continuously sampled (i.e., pulse-amplitude modulated) band-limited signal. Nevertheless, the method cannot be extended to discrete time sampling since it is based on the fact that the non-zero portions of the sampled waveform essentially comprise an infinite number of discrete samples. This can be stated formally in terms of function analyticity. Swaminathan [3] has used linear system identification techniques for signal restoration from data aliased in time. Since the method consists of modelling the causal and time-reversed anti-causal parts of the time-aliased signal as the impulse responses of stable, causal filters, it too cannot be used for the problem at hand. Powell [4] has enumerated the conditions under which a broad-band sparse spectrum is not destroyed by undersampling. However, only the particular band about the origin is protected from aliasing, and therefore the method cannot be applied to periodic signals, all of whose harmonics must be recoverable.

The only previously known practical algorithm for de-aliasing an undersampled periodic waveform has been given by Rader [1]. The Rader algorithm exploits the fact that samples obtained from many periods of a waveform can be sorted into a single period to dramatically increase temporal resolution, effectively removing aliasing distortion. While the same approach is used in conventional sampling oscilloscopes, these devices require operator intervention to adjust the triggering system so that the displayed periods truly correspond to the original waveform. The operator in effect must determine[2] when the proper signal period is being used to sort the samples, thereby relieving the oscilloscope of the most difficult task.

In both the Rader algorithm and the new algorithm to be presented in Chapter 4, the principal issue will be the determination of a signal's period. In both algorithms, waveform reconstruction is relatively straightforward once this has been accomplished. We will discuss the Rader algorithm in detail in Chapter 3, then implement and modify it. It will also serve as the basis for much of the other work in this thesis, the remainder of which is original for the most part.

[2]Or else provide a trigger signal whose period is the same as the waveform to be observed.

1.3 Scope, Contribution, and Organization of the Thesis

In Chapter 2, we will address theoretical issues which arise in sampling periodic waveforms. A new de-aliasing procedure and a new sampling criterion, both specifically for periodic signals, will be developed. Though stated for non-realizable conditions,[3] the new criterion will illustrate the upper bounds on performance which can be expected from the algorithms described in the chapters that follow.

The next two chapters contain detailed descriptions of algorithms for reconstruction of undersampled periodic waveforms. Chapter 3 describes the time domain de-aliasing algorithm mentioned briefly in the previous section. All work in Sections 3.1 and 3.2 is directly attributable to Rader [1,5], though some liberties have been taken in interpretation. Section 3.3 contains a new, simple modification of the relatively complex Rader algorithm, intended to increase algorithm efficiency when possible. Chapter 4 describes an original algorithm for de-aliasing in the frequency domain which, though perhaps not as elegant as the Rader algorithm, will be shown to be comparable in many instances. Typical reconstructions for natural and synthetic signals, along with other pertinent data, are presented at the conclusion of each of these two chapters.

The research is summarized in Chapter 5, in which we discuss the relative strengths and weaknesses of all algorithms and their variants, and perform empirical comparisons, as well. Issues such as speed, robustness, and reconstruction quality are considered. Suggestions for future research are enumerated in Chapter 6.

It will be most convenient to introduce new notation as it is needed. Whenever possible, results from previous works not directly related to de-aliasing will merely be stated, and appropriate references will be cited.

---

[3] A property it shares with perhaps all other criteria, including the Nyquist criterion.

Chapter 2

Development of a Sampling Criterion for Periodic Signals

The principal concern of this thesis is the recovery of a periodic continuous-time signal, of unknown frequency, from a set of uniformly spaced samples obtained using a sampling frequency below the Nyquist rate. This chapter will provide the necessary theoretical background, and more importantly, new extensions of conventional theory better suited for the problem at hand. Implementation issues will be treated in the chapters that follow.

A sampling criterion will be needed to indicate when a set of samples retain all of the information in a periodic analog signal. We will briefly review the classic Nyquist sampling criterion for lowpass and bandpass signals. The greater portion of the chapter will be devoted to reformulating the Nyquist criterion for the special case of periodic waveforms. In the process, a theoretical procedure for de-aliasing such signals will also be developed. Finally, methods for reducing ambiguity problems exposed during the development of the new criterion will be discussed.

2.1 The Nyquist Sampling Criterion

Many practical signals are generated by physical processes and as such, can be regarded as approximately band-limited by neglecting the minute amount of energy at

frequencies above a judiciously chosen cutoff[1] Ω_o. If such a signal is sampled uniformly, then there exists a minimum sampling rate for which the original signal can still be completely recovered.

Enumeration of a sampling criterion which specifies this minimum rate has been attributed to several authors, including [6]: Nyquist, Shannon, Whittaker, and Kotel'nikov. Because it was first introduced by Nyquist in 1928, in the context of telegraph transmission theory, we hereafter will refer to it as the *Nyquist sampling criterion*. The Nyquist criterion is well documented in the literature of signal processing and communications, as well as that of several other fields. It is repeated here only for completeness:

Criterion 2.1 *If an analog signal $x_a(t)$ contains no energy at frequencies Ω outside of the range $|\Omega| < \Omega_o$ rad/sec, then it is completely determined by its ordinates at a series of points equally spaced by π/Ω_o seconds or less.*

The Nyquist criterion actually applies to a wider variety of signals than just those of a lowpass nature. *Destructive aliasing will not occur in sampling any analytic*[2] bandpass signal which contains no energy outside of some range

$$-\Omega_o + \Omega_c \leq \Omega < \Omega_o + \Omega_c$$

provided that the sampling rate Ω_s is greater than or equal to the Nyquist rate $2\Omega_o$. In addition, a non-analytic signal sampled at Ω_s will not be aliased[3] if it contains no energy outside of the union of the ranges

$$-\frac{p+1}{2}\Omega_s \leq \Omega < -\frac{p}{2}\Omega_s$$

$$\frac{p}{2}\Omega_s \leq \Omega < \frac{p+1}{2}\Omega_s$$

where p is any integer. For each case above, if the respective parameter Ω_c or p is known, then the recovery procedure will be well defined. These are perhaps the two simplest cases to which the Nyquist criterion can be extended.

[1] The uppercase Ω will be used hereafter to denote continuous-time frequency (in radians/second), with the lowercase ω being reserved for the discrete-time case (radians/sample).

[2] One which has no energy at negative frequencies.

[3] We will use the term aliasing to imply *destructive* aliasing when clear from context.

The Nyquist criterion specifies a set of conditions which is sufficient but not necessary to permit reconstruction of a band-limited waveform from uniformly-spaced samples. Clearly, we can choose other criteria which may be more amenable to other signal representations, sampling methods, etc. In the next section, it will prove advantageous to do so, though we still must be sensitive to the basic issues of spectral overlap and reconstruction ambiguity.

2.2 The Pseudo-Nyquist Sampling Criterion

Many types of waveforms can be recovered from their samples even when they occupy a frequency band larger than the maximum permitted by the Nyquist criterion. Signals having sparse spectra form one such class, and include periodic signals and frequency-modulated narrow-band signals. They typically are non-analytic functions which do not meet the generalized Nyquist criterion passband requirement specified at the conclusion of section 2.1. Non-destructive undersampling of modulated signals is treated in [4], and will not be discussed further here. In this section, we will determine a set of conditions for which an undersampled periodic waveform can still be completely recovered, and outline a hypothetical reconstruction procedure. These conditions will then be incorporated in a new sampling criterion specifically for periodic signals.

We begin by addressing the issue of spectral overlap due to aliasing. An analog signal $x_a(t)$, periodic for all time, is characterized by a line spectrum $X_a(j\Omega)$. Sampling the signal over all time at constant intervals T yields a spectrum $X(e^{j\Omega T})$ exhibiting no spectral overlap unless two or more harmonics, each inherently having zero width, are aliased to the same frequency. Because $X(e^{j\Omega T})$ is periodic in Ω, we need only determine where all aliased harmonics appear in the baseband $0 \leq \Omega \leq 2\pi/T$ rad/sec in order to check for overlap.

Using a sampling rate[4] Ω_s, the n^{th} harmonic of a waveform with a fundamental frequency Ω_w is modulated down to $\langle n\Omega_w \rangle_{\Omega_s}$, where $\langle x \rangle_y$ denotes the quantity x modulo y. If the ratio Ω_s/Ω_w can be expressed as a rational number u/v with u and v in lowest terms (i.e., their *greatest common denominator* $(u, v) = 1$), then each harmonic

is aliased to one of only u (or fewer) distinct frequencies. The new location of the $(n-u)^{th}$ harmonic is

$$\langle (n+u)\Omega_w \rangle_{\Omega_s} = \langle n\Omega_w - v\Omega_s \rangle_{\Omega_s} = \langle n\Omega_w \rangle_{\Omega_s}$$

i.e., the same as the n^{th} harmonic.

If there are more than u consecutive analog harmonics, signal recovery is impossible since at least two harmonics overlap in the aliased spectrum. Therefore, unless a signal is known to contain fewer than u harmonics, we must require Ω_s/Ω_w to be irrational. If the latter condition is met, all harmonics will be aliased to unique frequencies. The number of harmonics must be finite, but is otherwise unrestricted and in fact can be unknown.

In order to determine Ω_w, we first must be able to identify the locations[5] of the aliased harmonics. If the number of non-zero harmonics is finite, then their locations can be detected and stored in a list. If the first harmonic in the analog waveform is non-zero, its location after aliasing ($\langle \Omega_w \rangle_{\Omega_s}$) will be included in the list above. We only need to determine the list entry to which it corresponds.

Suppose the periodic analog signal is band-limited to *any* known range, $|\Omega| \leq \Omega_h$.[6] This clearly guarantees a finite number of harmonics. Another requirement is needed: either the (analog) spectral component at Ω_w (which we will call "Ω_1"), the component at $-\Omega_w$ ("Ω_{-1}"), or both must be non-zero. For the common case of real signals, we must require that both be non-zero. We do not need to know which of the cases above is true, but at least one of the harmonics at Ω_1 and Ω_{-1} is necessary in the recovery procedure to follow in order to determine Ω_w.

The first step of the procedure consists of listing the abscissas, i.e., frequency locations, of all spectral lines in the region $0 < \Omega < \Omega_s$. Each value is then used as a guess

[4] In general, subscripts s will denote quantities related to the sampler, and subscripts w will correspond to the waveform to be reconstructed.

[5] At least in theory, harmonic amplitudes do not help in determining Ω_w, only in the subsequent reconstruction process.

[6] Note that the choice of Ω_h is completely arbitrary, viz., independent of both Ω_w and Ω_s. Therefore, cases in which $\Omega_h \gg \Omega_s$ are acceptable.

of the aliased fundamental, $\langle \Omega_w \rangle_{\Omega_s}$. We observe the spectrum at positive and negative multiples (interpreted modulo Ω_s) of each guess. The arbitrary but known signal cutoff Ω_h indicates when to stop this process in each direction along the Ω-axis. The number of multiples for which the spectrum is non-zero is recorded.

The guess yielding the maximum tally must be either $\langle \Omega_1 \rangle_{\Omega_s}$ or $\langle \Omega_{-1} \rangle_{\Omega_s}$ (the latter $= \langle -\Omega_1 \rangle_{\Omega_s}$). If both are present, there will be two "best" guesses. Each incorrect guess, corresponding to an analog frequency Ω_N or Ω_{-N} (the harmonics at $N\Omega_w$ and $-N\Omega_w$, respectively, where $N > 1$), results in a lower tally because only one of every N harmonics which might be non-zero has been counted. The numbers of positive and negative harmonics in the waveform need not be equal. In addition, missing harmonics cause no harm unless both "Ω_1" and "Ω_{-1}" are absent.

We now must itemize any additional contraints which are mandatory for obtaining Ω_w unambiguously from the value(s) found above. The maximum unambiguous range of Ω_w cannot be greater than or equal to Ω_s. Consider two signals with the same waveshape (or equivalently, the same Fourier series coefficients) but different fundamental frequencies Ω_A and $\Omega_B = \Omega_A + r\Omega_s$, where r is some integer. The n^{th} harmonic from each is aliased to the same frequency, rendering the two sampled signals indistinguishable.

Unfortunately, the restriction above is insufficient. Whether real or complex, a periodic signal might have energy at both positive and negative multiples of its fundamental. Since

$$\langle -n\Omega_w \rangle_{\Omega_s} = \langle \Omega_s - \langle n\Omega_w \rangle_{\Omega_s} \rangle_{\Omega_s}$$

the $-n^{th}$ and n^{th} harmonics will be aliased to mirror image locations about $\Omega_s/2$. In listing aliased-harmonic abscissas as done above, the same set of entries are obtained from a signal of frequency Ω_A and another of frequency $\Omega_B = -\Omega_A + r\Omega_s$, where r is any integer, if the same[8] harmonics are present. If $r = 0$ and the two signals have the same harmonic coefficients, one signal is simply the time reversal of the other. We must know that Ω_w lies in a particular range $p\Omega_s/2$ to $(p+1)\Omega_s/2$, for some integer p. The maximum unambiguous range is thus only $\Omega_s/2$, *and* it cannot span any multiple

[8]This refers to the harmonic numbers ($\pm 1^{st}$, $\pm 2^{nd}$, ...) and does not concern the harmonic amplitudes.

of $\Omega_s/2$.

If there is only one best guess Ω_{best}, the value of p indicates whether or not to negate it. If there are two best guesses, the value of p uniquely determines which of the two to use since Ω_w can only differ from $\langle\Omega_w\rangle_{\Omega_s}$ by a multiple of Ω_s. (p indicates the proper frequency range of width $\Omega_s/2$.) After negation (if necessary), the appropriate multiple of Ω_s is then added, and we proceed to reconstruction. The latter consists of unravelling the aliased harmonics, and is simple once Ω_w and Ω_s are both known. Figure 2.1 contains a flowchart summarizing the procedure described above. It is assumed that Ω_s/Ω_w is irrational, and that Ω_s, Ω_h, and p are known.

There exists at least one inefficiency in the de-aliasing procedure described above. We checked for non-zero harmonics at positive and negative multiples of $\langle\Omega_N\rangle_{\Omega_s}$, not Ω_N. Because these multiples were interpreted modulo Ω_s as well, the exact same sequence of spectral locations would have been checked if we had known and used Ω_N and its multiples instead. The only difference concerns just how quickly the process would have terminated in each of the positive and negative frequency directions.

Had we used Ω_N, we properly would have stopped searching the aliased spectrum when $n\Omega_N > \Omega_h$. However, the termination condition we actually used was $n\langle\Omega_N\rangle_{\Omega_s} > \Omega_h$. We effectively checked for analog harmonics above Ω_h. Since these harmonics were non-existent, the tallies remained undistorted. If the Ω_w range parameter p was known, we could have adjusted each guess $\langle\Omega_N\rangle_{\Omega_s}$ beforehand to lie in the allowable range. However, the generality gained from not requiring this will be advantageous later.

Based on the above discussion, we now define a new sampling criterion for periodic signals which we will call the *pseudo-Nyquist sampling criterion*:

Criterion 2.2 *If an analog signal $x_a(t)$ is periodic, contains no energy at frequencies Ω outside any range $|\Omega| < \Omega_h$ rad/sec, and its fundamental frequency Ω_w lies in the range $p\Omega_o/2$ to $(p+1)\Omega_o/2$ where p is an integer and the quantity Ω_o/Ω_w is irrational, then it is completely determined by its ordinates at series of points spaced apart by $2\pi/\Omega_o$ seconds.*

We would be able to relax one limitation imposed by the pseudo-Nyquist criterion

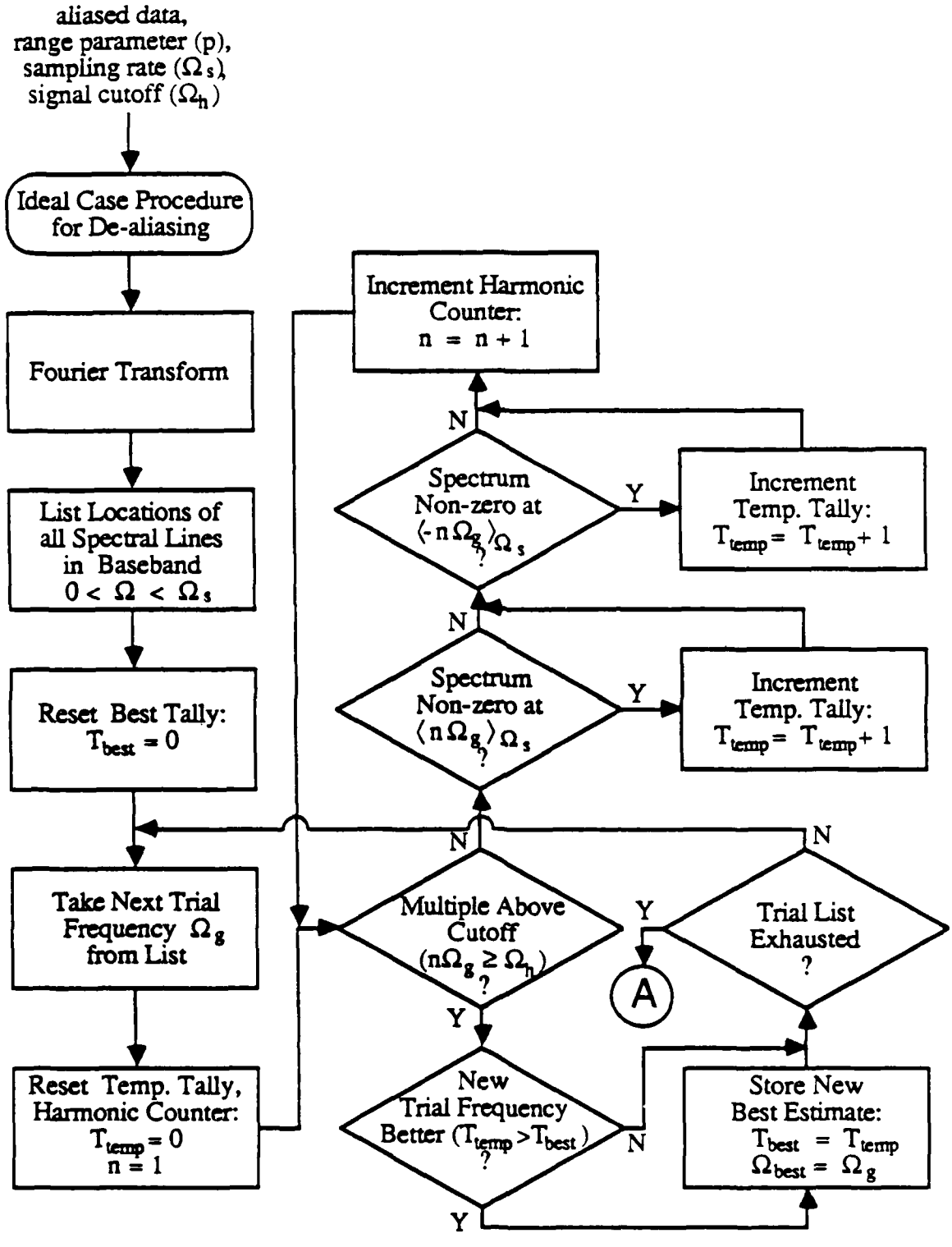
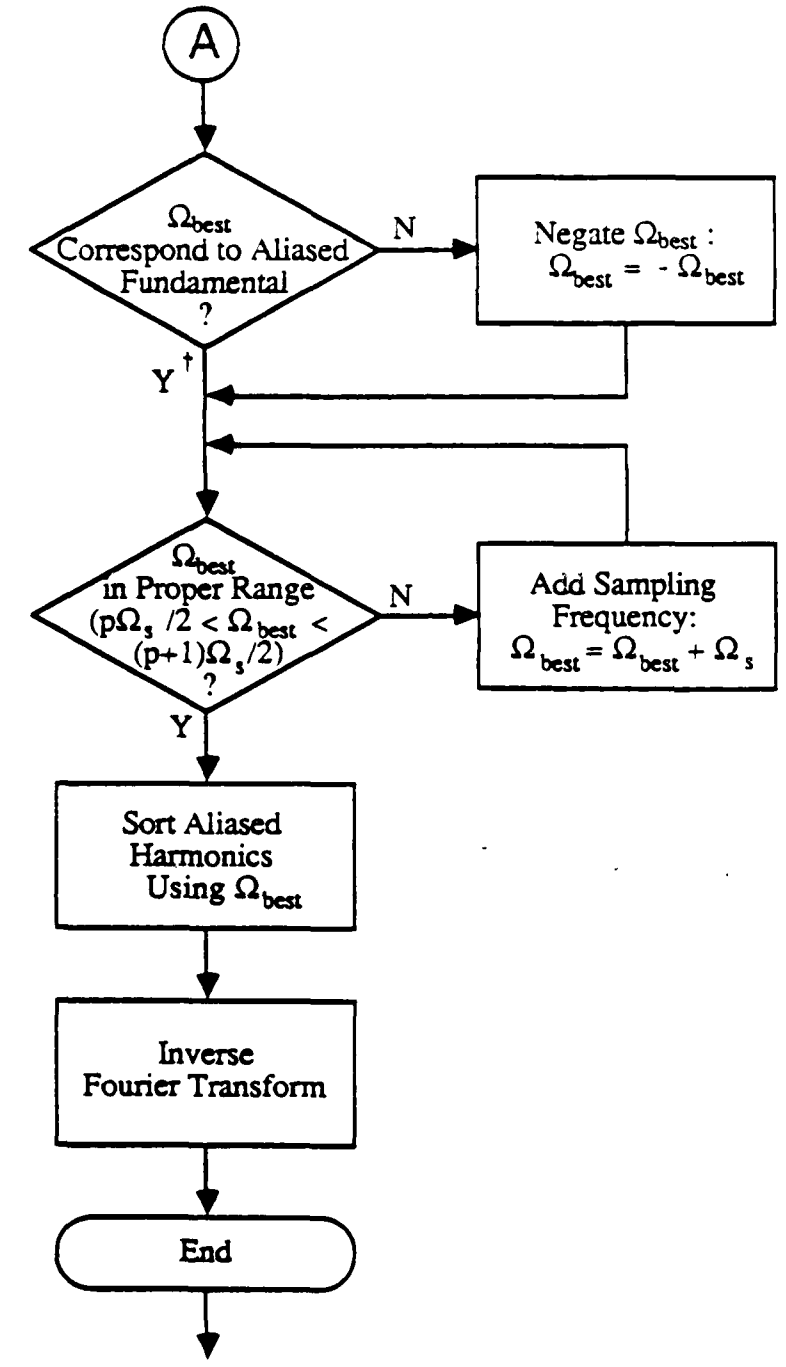


Figure 2.1: Ideal case procedure for recovering an aliased signal in the frequency domain.



Notes:
† Equivalently, $\Omega_{best} < \Omega_s/2$ if p even, $\Omega_{best} > \Omega_s/2$ if p odd.

Figure 2.1: *continued*

if it were not for the fact that the number of harmonics is unknown. Since only a finite number of harmonics can be present, Ω_s/Ω_w need not be irrational. Recall that if Ω_s/Ω_w can be expressed as a rational number u/v where $(u,v) = 1$, then up to u consecutive harmonics can still be present in the signal without resulting in destructive aliasing. The probability of this happening with u simultaneously being prohibitively small is very low.

Finally, consider the effects of p being unknown. In this case, we could temporarily assume $p = 0$. Determination of Ω_w and signal recovery would proceed in *exactly* the same manner as before. The only difference(s) between the reconstructed and true waveforms would be a constant scale change along the time axis and/or a reversal of time. There are probably applications where this is tolerable. If if is not, there are still means for effectively removing these two ambiguities, as described in the next section.

2.3 Reducing Signal Fundamental Frequency Ambiguity

Two ambiguities arise when the range of permissible values for the signal fundamental frequency Ω_w is unknown. (We will assume this is true for the remainder of the chapter.) Given only the sampling frequency Ω_s and setting $p = 0$, the procedure from the previous section yields a unique value in the range 0 to $\Omega_s/2$ corresponding to either $\langle\Omega_w\rangle_{\Omega_s}$ or $\langle-\Omega_w\rangle_{\Omega_s}$. We cannot determine which of the two it is, and even if we could, we would not know what multiple of Ω_s to add to that value (negated if necessary) in order to obtain Ω_w.

Two possible solutions to the first problem are:

1. Only allow periodic waveforms which are analytic.

2. Filter non-analytic signals with a Hilbert transformer, before sampling, to remove all energy at negative frequencies.

Either one insures that the ambiguous value found is identically $\langle\Omega_w\rangle_{\Omega_s}$ since the analog harmonic at $-\Omega_w$ has zero amplitude. If a signal is not analytic and waveform

reconstruction is required (in addition to a value for Ω_w), then both the filtered and original signals must be sampled. Samples of the former are needed for determining Ω_w, and those of the latter for signal recovery.

The *Chinese remainder theorem* from number theory provides a convenient solution to the second ambiguity problem mentioned previously. Before describing this we first present some necessary notation from number theory. $(x)_M$ denotes the *residue* of x modulo the *modulus* M. This residue is defined as the remainder of x divided by M. Since all integers $x + kM$ (for arbitrary k) are *congruent*, i.e., they yield the same residue modulo M, they are said to form a *residue class* modulo M. There are M residue classes. Using this notation, the Chinese remainder theorem can be stated as follows:

Theorem 2.1 *The congruences $(x)_{m_i} = r_i$ possess a unique solution among the residue classes modulo $M = \prod m_i$ if the moduli m_i are mutually prime in pairs. The solution for x is the residue class $R = \sum r_i N_i M_i$ where each $M_i = M/m_i$, and each N_i is the solution of an equation $(N_i M_i)_{m_i} = 1$.*

In the above theorem, all variables are integers. Proofs of the theorem can be found in most texts on number theory [7,8,9,10].

Using the Chinese remainder theorem, several highly ambiguous residues r_i of an unknown quantity x can be combined into a single, much less ambiguous residue R, provided that the moduli m_i are pairwise coprime. The uncertainty range of each of the residues r_i is the corresponding modulus m_i, while the uncertainty range of R is $\prod m_i$.

Application of the Chinese remainder theorem is not restricted to problems involving only integers, however. It can be utilized for rational operands, as well. Since all practical situations involve finite precision arithmetic, all quantities are rational, regardless of the units used. Given the units and the size of a quantum, we first normalize the dimension of interest in terms of a unit quantum. Integral,[9] mutually prime moduli are then chosen, and integral residues are found. The single unambiguous (or at least

[9] After normalization.

less ambiguous) residue determined using the Chinese remainder is then de-normalized to yield the desired quantity.

Suppose that after normalizing time in the manner above, we sample an analytic periodic signal at several integral, mutually prime sampling rates, simultaneously. The procedure in Section 2.2 can be used to produce a residue (viz., the frequency of the aliased fundamental harmonic) from each of the resulting sequences. If the residues from this ideal procedure are quantized, a unique value of the true fundamental frequency modulo the *product* of the sampling rates can be obtained. By using either higher sampling rates or, more appropriately from our standpoint, additional sampling systems, the ambiguity problem can be virtually eliminated.

Consider the following simple example. The clock rates of four samplers are 7, 8, 9, and 11 samples per second, respectively. All measurements are to be quantized in Hertz. Using the output from each of the four samplers in the procedure from the previous section, we obtain values for the aliased fundamental frequency of 2, 5, 5, and 6 Hz, respectively. To get the true value of the fundamental frequency f_w, we utilize the Chinese remainder theorem: x is f_w; the sampling rates are $m_1 = 7$, $m_2 = 8$, $m_3 = 9$, and $m_4 = 11$; and the residues are $r_1 = 2$, $r_2 = 5$, $r_3 = 5$, and $r_4 = 6$. Therefore,

$$\begin{aligned} M &= 7 \cdot 8 \cdot 9 \cdot 11 = 5544 \\ M_1 &= 8 \cdot 9 \cdot 11 = 792 \\ M_2 &= 7 \cdot 9 \cdot 11 = 693 \\ M_3 &= 7 \cdot 8 \cdot 11 = 616 \\ M_4 &= 7 \cdot 8 \cdot 9 = 504 \end{aligned}$$

Continued fractions [10] can be used to solve

$$\begin{aligned} \langle 792 N_1 \rangle_7 &= 1 \\ \langle 693 N_2 \rangle_8 &= 1 \\ \langle 616 N_3 \rangle_9 &= 1 \\ \langle 504 N_4 \rangle_{11} &= 1 \end{aligned}$$

yielding $N_1 = 1$, $N_2 = 5$, $N_3 = 7$, and $N_4 = 5$. Finally,

$$R = r_1(792 \cdot 1) + r_2(693 \cdot 5) + r_3(616 \cdot 7) + r_4(504 \cdot 5)$$

The values N_i can be pre-computed and reused for any set of measurements r_i. Entering the present values of r_i into the formula above yields $R = 149$. This is the residue class modulo 5544 Hz to which the fundamental frequency belongs. Equivalently, $f_w = 149 + 5544j$ Hz, for some unknown integer j. If f_w is known to lie in some range whose width is less than or equal to 5544 Hz, then it can be uniquely determined from the four aliased sequences above. If a greater unambiguous range is desired, one or more additional samplers with appropriate clock rates will be required.

The usefulness of the Chinese remainder theorem is readily apparent from the example above. For any one sampler used alone, the maximum unambiguous range of f_w would have been the sampling rate, less than 12 Hz. But because the four clock rates are pairwise mutually prime, the maximum unambiguous range was extended to greater than 5 kHz.

Chapter 3

Rader Time Domain Sample Sorting Algorithm

In the previous chapter, we specified a set of conditions under which a periodic signal can be completely recovered from its samples, even after undersampling. However, these conditions cannot be met, and therefore much of the remainder of this thesis will be devoted to the practical aspects of the recovery problem.

In this chapter we will review the theory and discuss our implementation of an efficient time domain de-aliasing algorithm developed by Rader [1]. An iterative technique is used for determination of the signal period T_w, and constitutes the bulk of the processing required. Subsequent waveshape recovery consists of time series sorting, and is straightforward once T_w is known. Results from number theory are exploited to make the approach practical.

Section 3.1 will describe the general approach of the Rader algorithm. It will include the development of a criterion proposed by Rader for indicating the best reconstructed signal among several trial reconstructions, simultaneously providing an estimate of T_w. The second section will discuss the algorithm in detail, and will include flowcharts summarizing our implementation of it. Unless noted otherwise, all work to be described in Sections 3.1 and 3.2 is due to Rader [1,5], though some liberties will be taken in interpretation. In a few instances, it will be beneficial to supplement the discussions provided by Rader.

Section 3.3 will discuss a new modification of the Rader algorithm in which the iterative procedure for estimating T_w is accelerated by decomposing it into a series of successively finer searches, with a coarse search being used on the first iteration. Typical reconstructions for undersampled natural and synthetic signals will be presented in the closing section, along with other pertinent data.

It will be convenient to normalize time using the sampling period T_s. We will refer to the ratio T_w/T_s as τ_w, the normalized waveform period (or simply the waveform period, when clear from context). More generally, τ $(= t/T_s)$ will be used as a dimensionless independent variable for continuous time. Likewise, we will define a normalized frequency variable[1] $\phi = \Omega/\Omega_s$, where $\Omega_s = 2\pi/T_s$.

In both this chapter and the following one, we shall assume that T_s is known, and that T_w is not. Since all processing involves time-normalized data, the same algorithms can be used when the reverse is true. The degree of accuracy to which T_s is known will not be critical in any of the algorithms presented in this thesis, since the value is only needed for computing the output sample spacing. For now, we will also assume that both T_w and T_s are stable. The repercussions of unstable periods will be discussed in Chapter 5.

3.1 General Approach

If both the signal and sampling periods are known, waveform recovery is simple. Each sample $x[n]$, corresponding to the analog signal $x_a(t)$ at $t = nT_s$, is equal to the sample that would have been obtained at time $t = \langle nT_s \rangle_{T_w}$. To recover the original waveshape, we can place each sample in a *composite period* at $t = \langle nT_s \rangle_{T_w}$, or equivalently, $\tau = \langle n \rangle_{\tau_w}$. The composite period thus extends over the range $0 \leq \tau < \tau_w$. We have chosen to view its formation as wrapping the samples onto a cylinder of circumference τ_w, as depicted in Figure 3.1.

The sample spacing within a composite period typically is not uniform. In general,

[1] We use ϕ, measured in revolutions, to distinguish it from f, Ω and ω, typically corresponding to quantities measured in Hertz, radians/second, and radians, respectively.

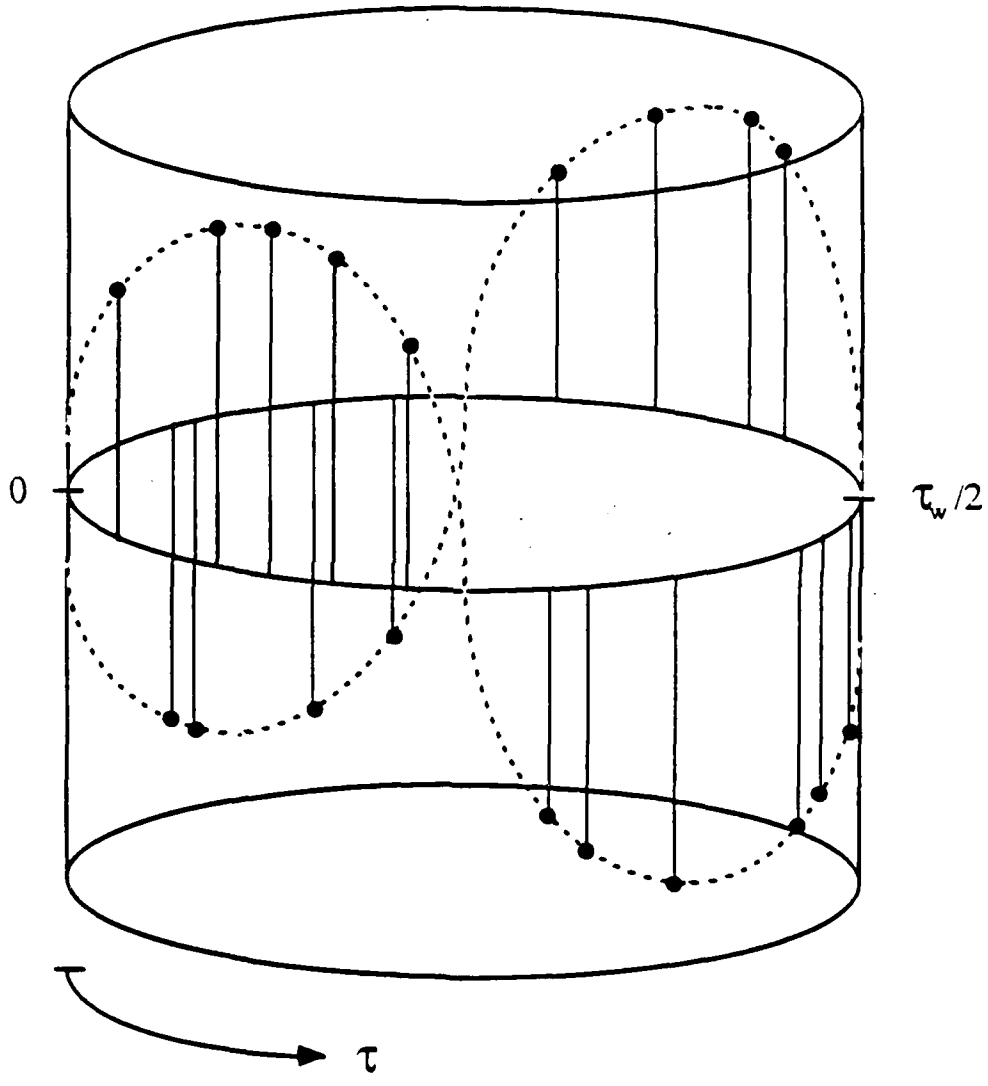


Figure 3.1: Formation of a composite period.

successive samples $x[n]$ are scattered to non-integral locations along the τ-axis. Nevertheless, the waveshape of the original signal $x_a(t)$ should be apparent, even for modest quantities of samples and for arbitrarily large T_s.

We know from Chapter 2 that the procedure above generally will fail if τ_w is rational. In fact, the irrationality requirement in the pseudo-Nyquist criterion can be justified with a time domain argument similar to the frequency domain argument presented earlier. If τ_w ($= \Omega_s/\Omega_w$, from Chapter 2) can be expressed as a rational number u/v where $(u, v) = 1$, then each sample has one of only u (or fewer[2]) distinct ordinates. Since

$$\langle n + u \rangle_{\tau_w} = \langle n + v\tau_w \rangle_{\tau_w} = \langle n \rangle_{\tau_w}$$

the $(n + u)^{th}$ sample is identical to the n^{th} sample.

However, we also know from Chapter 2 that if τ_w can be expressed as a fraction u/v as above, destructive aliasing still will not occur if all harmonics in the original waveform spectrum occupy u or fewer spectrally adjacent harmonic locations. If this is the case, the signal can be completely recovered by taking only the first u (i.e., a unique subset) of N available samples (assuming $N \geq u$), forming a composite period, and interpolating as desired. The interpolation method must be insensitive to non-uniform sample spacing.

The sample sorting algorithm above is insufficient for the more common case where the signal period is *not* known. However, suppose that we repeat the reconstruction process for several guessed or *trial* periods τ_g, one of which is the correct period, τ_w. Assuming enough samples are used, it is not unreasonable to expect the composite period formed with the true period to be "smoother" than the others so formed.

In order to implement such an iterative technique, we need a method for estimating the "smoothness" of a composite period. For this purpose, Rader has defined the *variation* of a composite period as the sum of the absolute values of the differences between successive composite period samples $x_{\tau_g}[n]$, including the last and first samples.[3] The

[2]In the case where the ordinates of one period of the analog waveform are not unique.

subscript τ_g indicates the trial period used to form the composite period.

$$\mathcal{V}(\tau_g) = x_{\tau_g}[0] - x_{\tau_g}[L-1] - \sum_{n=1}^{L-1} x_{\tau_g}[n] - x_{\tau_g}[n-1] \tag{3.1}$$

L is the number of samples available, and for now, also the number in the composite period. The indices of the $x_{\tau_g}[n]$ only indicate temporal ordering. They do not imply uniform sample spacing.

If we were to reconstruct an aliased sinewave with amplitude A using its true period τ_w and many samples (so that the composite period contained samples near the maximum and the minimum of the sinewave period), $\mathcal{V}(\tau_g)|_{\tau_w}$ would be very nearly, if not exactly, $4A$. The sinewave samples would be in the wrong temporal order if an incorrect period was used. This would yield a larger value of $\mathcal{V}(\tau_g)|_{\tau_w}$, unless $1/\tau_g$ and either $1/\tau_w$ or $-1/\tau_w$ were congruent modulo one ("$1/\tau_s$", where τ_s is the normalized sampling frequency), in which case $\mathcal{V}(\tau_g)|_{\tau_w}$ would be the same. (Refer to Section 2.2.) We would expect similar results for many other types of waveforms, including those rich in high frequency components.

Based on the assumptions above, Rader has proposed the following criterion for choosing the "best" value of τ_g from a properly chosen, finite set used in the prescribed manner:

Criterion 3.1 *The trial period which yields the waveform of smallest variation is the correct period, and the resulting waveform is the correct waveform.*

We will refer to this as the *minimum variation criterion.* The choice of a suitable set of trial periods will be treated in Section 3.2.

It is probably impossible to justify the criterion deterministically. This might be made possible by redefining variation using squares rather than absolute values of successive differences. Since the criterion has yet to be proven using either definition, the original one should be retained for a purely practical reason. Most of the processing required by the algorithm described in the next section involves computation of many

[3]The bracket notation used for the time variable n is somewhat misleading since, in the most general sense, x_{τ_g} is a function of a continuous variable (τ). However, it will be accurately described as a discrete time function when implemented.

trial variations. Therefore, using squares (viz., multiplies) instead of absolute values would incur a substantial penalty.

The minimum variation criterion can be supported, however, with a probabilistic argument. Although the manner in which we state the argument here is different from that used by Rader, the key issues remain unchanged. We will consider the effect of using a given trial period τ_g with increasing N, the quantities of samples used. Two cases will be examined: $\tau_g = \tau_w$, and $\tau_g \neq \tau_w$. In either case, as more samples are used, the variation may increase, and it cannot decrease. However, the effects in the limit (as $N \rightarrow \infty$) are distinct in each case.

Suppose that we form several composite periods using the correct period τ_w (which must be irrational) with different N. As N increases, the variation[4] $\mathcal{V}_N(\tau_w)$ asymptotically approaches $\mathcal{V}_a$, the "variation" of one period of the original analog signal:

$$\lim_{N \rightarrow \infty} \mathcal{V}_N(\tau_w) = \mathcal{V}_a \tag{3.2}$$

In the limit, there would be no inflections (local maxima or minima) of the original waveshape between any two successive samples in the reconstructed period. An example involving four different values of N is shown in Figure 3.2. Note that each of the variations for the last three plots is approximately equal to 16 (i.e., $\mathcal{V}_a$).

Now suppose that an incorrect period τ_g ($\neq \tau_w$) is used. Increasing N should always result in a larger variation. $\mathcal{V}(\tau_g)$ almost certainly will increase without bound since, in the limit, each ordinate of the original waveform will be next to every other, after formation of the composite period. It thus seems reasonable that the minimum variation criterion will hold for finite N when N is somewhat greater than the number of significant harmonics present in the waveform, since the latter governs the number of inflections in a true period of the original waveform. Empirical evidence (viz., plots of actual variation functions for various N) will be presented in Section 3.4, along with all other experimental results pertaining to this chapter.

We can now discuss the algorithm provided by Rader to implement the preceding procedures for determining τ_w and recovering the original waveform $x_a(t)$.

[4] As defined using absolute, not squared, differences.

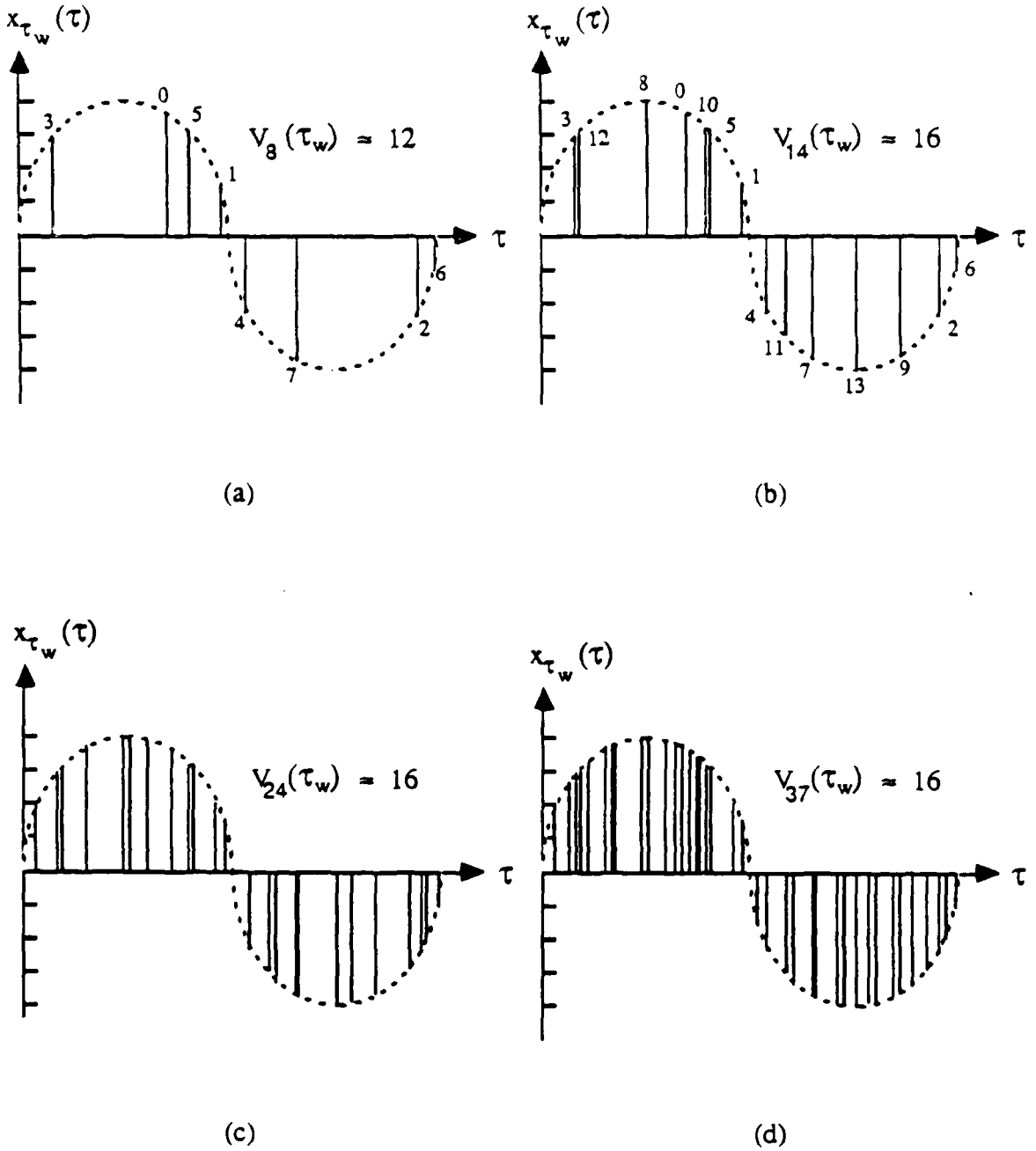


Figure 3.2: Composite periods formed using correct trial period ($\tau_g = \tau_w$) and various quantities of samples. Numbers on (a) and (b) correspond to indices n of original aliased sequence $x[n]$.

3.2 Detailed Description of the Algorithm

Once we have $\mathcal{V}(\tau_g)$, the composite period variation for all τ_g, we should be able to determine the true signal period τ_w (and thus recover the original waveform) using the minimum variation criterion presented in the preceding section. However, two problems must be circumvented in computing $\mathcal{V}(\tau_g)$: it is a function of a continuous variable τ_g, and it has infinite extent in this dimension. We are limited to finite search ranges for τ_g, and only those composed of discrete points.

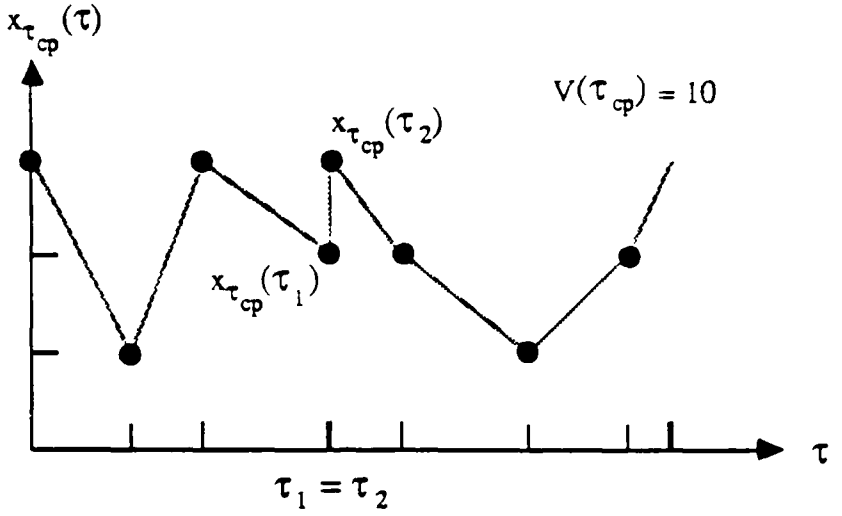
In developing the pseudo-Nyquist criterion we showed that sampling two signals having identical Fourier coefficients yields identical sequences $x_1[n]$ and $x_2[n]$ if the signal fundamental frequencies differ by a multiple of the sampling rate. We also showed that if the sum of their fundamental frequencies is a multiple of the sampling rate, one sequence is the time reversal of the other. (The consequences of ignoring phase are minimal here.) Therefore, the limits on the search range, τ_{min} and τ_{max}, must be chosen such that their reciprocals do not span a multiple of $1/2$. This includes the requirement that

$$\frac{1}{\tau_{min}} - \frac{1}{\tau_{max}} < \frac{1}{2} \tag{3.3}$$

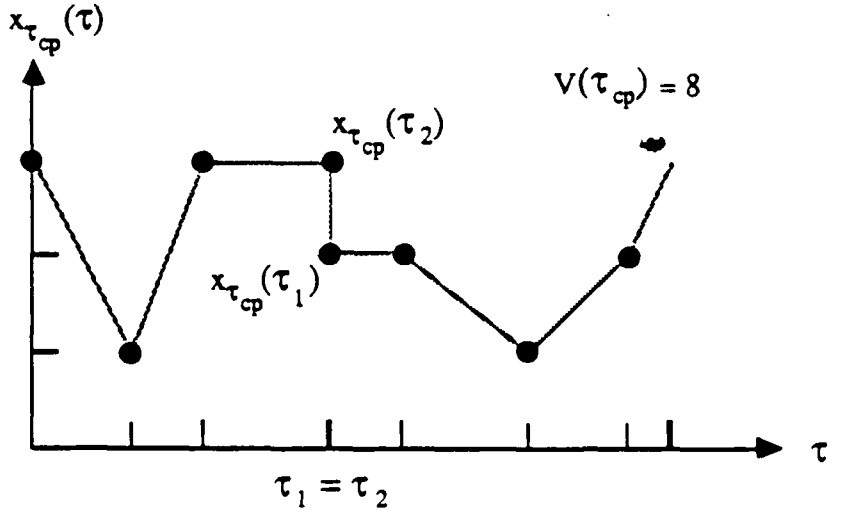
No additional restrictions need to be imposed in order to insure finite search ranges.

We now direct our attention to the need for a discrete set of trial periods. Fortunately, the function $\mathcal{V}(\tau_g)$ is always piecewise-constant. To show this, Rader first defined a *critical period* τ_{cp} as a value of τ_g for which two or more samples $(x_{\tau_g}(\tau_1), x_{\tau_g}(\tau_2), \ldots)$ in the corresponding composite period $x_{\tau_g}(\tau)$ would have the same abscissa $(\tau_1 = \tau_2 = \cdots)$, as shown in Figure 3.3. Referring back to Figure 3.1, we see that in continuously varying τ_g (which replaces τ_w as the circumference of the cylinder), the location of the n^{th} sample $x_{\tau_g}(\langle n \rangle_{\tau_g})$ also varies continuously. Note that $\mathcal{V}(\tau_g)$ cannot change unless two or more samples interchange. It is ambiguous at each critical period, and constant between any two which are adjacent.

A hypothetical variation function is shown in Figure 3.4. The limits of the search range, τ_{min} and τ_{max}, and the (unknown) true signal period τ_w are labelled. All other markers correspond to locations of critical periods.



(a)



(b)

Figure 3.3: Ambiguity of variation for composite period formed using a critical period ($\tau_g = \tau_{cp}$).

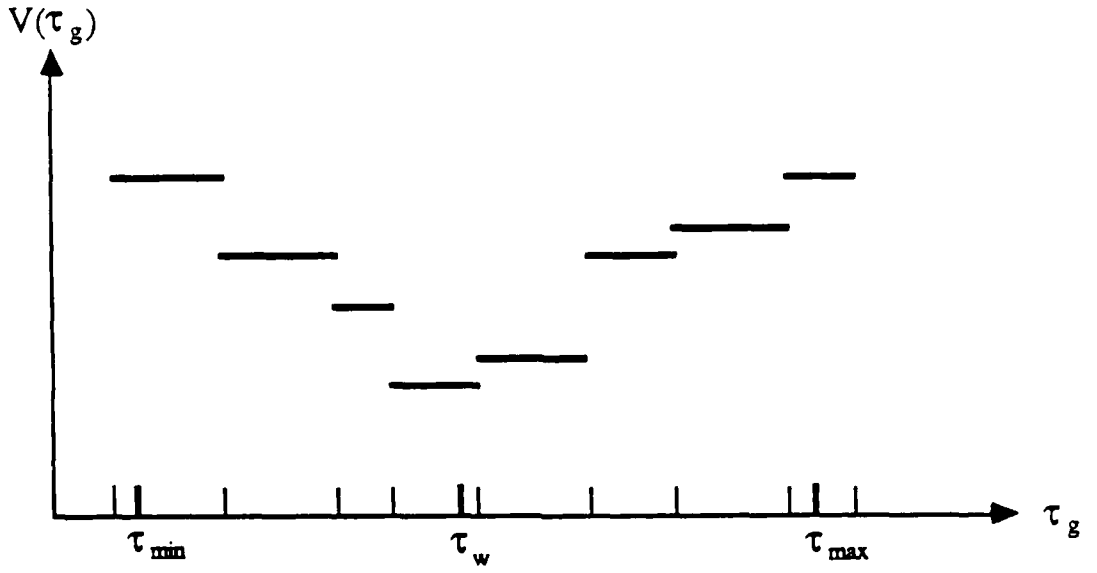


Figure 3.4: Hypothetical variation function showing ambiguity at critical periods (unlabelled markers).

We only need to compute $\mathcal{V}(\tau_g)$ at one point between each pair of adjacent critical periods. Any such value of τ_g can be used, though we will see that certain choices yield faster execution than others. According to the minimum variation criterion, the range of τ_g over which the variation is smallest should contain the true period τ_w. If we retain the value of τ_g in this region at which $\mathcal{V}(\tau_g)$ was computed as our estimate of τ_w, the corresponding composite period can be used as the reconstructed waveform. We will call this estimate τ_{best}. The samples in the composite period formed using τ_{best} will presumably have the same temporal ordering they would have had using the exact value of τ_w instead. Since τ_{best} cannot be a critical period, all samples will have unique abscissas.

Rader has shown that the critical periods τ_{cp} can be found by solving congruences relating the abscissas τ_1 and τ_2 of any two composite period samples $x_{\tau_g}(\tau_1)$ and $x_{\tau_g}(\tau_2)$ which would coincide ($\tau_1 = \tau_2$). If these two samples are the m^{th} and p^{th} samples from the unsorted sequence $x[n]$, then

$$\langle p \rangle_{\tau_{cp}} = \langle m \rangle_{\tau_{cp}} \tag{3.4}$$

Since we are only sorting a finite set of samples $x[n]$ where $n = 0, \ldots, N-1$, and the roles of the two samples are interchangeable, we can assume

$$0 < (p - m) < N \tag{3.5}$$

Defining the difference $p - m$ as j, we note that

$$\langle j \rangle_{\tau_{cp}} = 0 \tag{3.6}$$

or, equivalently

$$j = k\tau_{cp} \tag{3.7}$$

where k is a positive integer which cannot be greater than L, the maximum number of periods between the m^{th} and p^{th} samples:

$$k \leq L \tag{3.8}$$

where

$$L = \left\lfloor \frac{N-1}{\tau_{min}} \right\rfloor \tag{3.9}$$

The delimiters $\lfloor \rfloor$ indicate the integer part or "floor" function.

In summary, a critical period τ_{cp} can be expressed as a ratio

$$\tau_{cp} = \frac{j}{k} \tag{3.10}$$

where

$$0 < j < N \tag{3.11}$$

and

$$0 < k \leq L \tag{3.12}$$

Given a search range $[\tau_{min}, \tau_{max}]$ which satisfies the constraints enumerated in the pseudo-Nyquist criterion, we can list all possible critical periods satisfying Equations 3.11 and 3.12. However, if τ_{min} is small and/or N is large, the number of critical periods may be enormous. Sorting them (to determine which ones form adjacent pairs) would be an arduous task, and storage requirements could be prohibitive. An algorithm for generating successive critical periods is desirable.

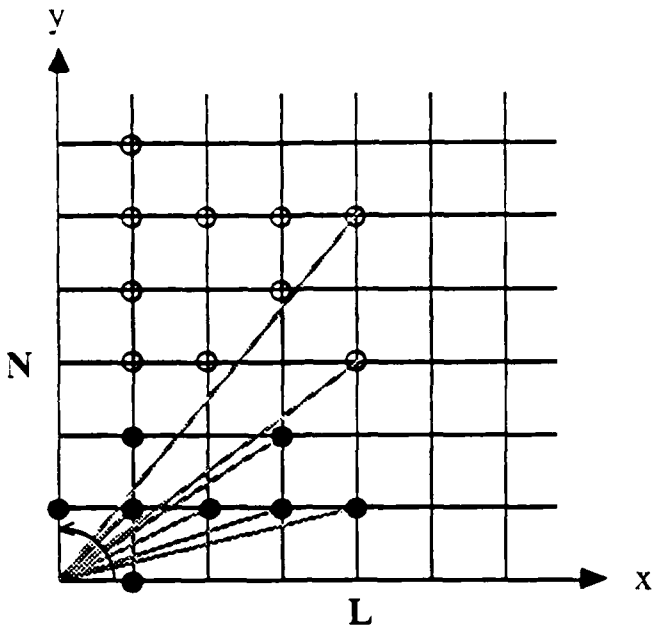


Figure 3.5: Generation of a Farey series of order $L = 4$, $\mathcal{F}_4 = \left\{\frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}, \frac{5}{4}, \frac{4}{3}, \ldots\right\}$. Solid dots indicate Farey fractions (y_o/x_o) which are also critical periods for $N = 3$.

Rader has indicated that the critical periods τ_{cp} form a sparse subset of a *Farey series* [7,8,9]. A Farey series of order L is defined as the sequence of all rational numbers u/v (where $(u, v) = 1$) whose denominators do not exceed L, arranged in increasing numerical order. For our purposes, the Farey fraction order is given by Equation 3.9. If the numerator of a Farey fraction is less than N (Equation 3.11), then it is also a critical period.

A graphical interpretation of the generation of a Farey series of order $L = 4$ is given in Figure 3.5. Posts are placed on a grid (perpendicular to it) at all integral locations in the first quadrant whose x-coordinates are less than or equal to L. An observer is placed at the origin of the grid, and is instructed to sweep his line of sight counterclockwise and name only the coordinates (x, y) of each post he can see. Each succeeding pair (x, y) forms the next member y/x of the Farey series. This series is indicated by the collective dots in Figure 3.5. Farey fractions which are also critical periods (for $N = 3$) are marked with solid dots.

Of course, a graphical method is not suitable for our purposes. However, given two

successive Farey fractions a/b and c/d of order L, it is possible to generate the next one, e/f. Specifically, let

$$Z = \left\lfloor \frac{L-b}{d} \right\rfloor \tag{3.13}$$

The next Farey fraction is then given by

$$\begin{aligned} e &= Zc - a \\ f &= Zd - b \end{aligned} \tag{3.14}$$

The proof is as follows. From [7] we know that two successive Farey fractions a/b and c/d of order L satisfy both

$$cb - ad = 1 \tag{3.15}$$

and

$$b + d > L \tag{3.16}$$

If e/f follows c/d in the Farey sequence, then $ed - cf$ must equal 1. Any $e = Zc - a$ and $f = Zd - b$ satisfies this equality. We therefore must find a value of Z satifying both

$$Zd - b \leq L \tag{3.17}$$

and

$$d + (Zd - b) > L \tag{3.18}$$

Equivalently

$$Z \leq \frac{L+b}{d} < Z + 1 \tag{3.19}$$

The unique solution is given by Equation 3.13.

Given two successive Farey fractions u_o/v_o and u_1/v_1 spanning τ_{min}, we can use Equations 3.13 and 3.14 to generate all the rest. We store the first Farey fraction u_o/v_o, then alternate between searching for the next critical period and computing $\mathcal{V}(\tau_g)$ for some τ_g between that critical period and the previous one. If a new $\mathcal{V}(\tau_g)$ is less than the previously stored minimum, it replaces that value, and the corresponding τ_g is also recorded.

We can use u_o/v_o for τ_g in the initial iteration if it is not a critical period. The last variation is computed when a critical period greater than τ_{max} is generated. Critical

periods are found by testing each new Farey fraction for a numerator less than N. Any Farey fractions between the previous and current critical periods ($\tau_{cp,prev}$ and $\tau_{cp,cur}$) are stored separately for use in computing the variation.

Recall that any value between $\tau_{cp,prev}$ and $\tau_{cp,cur}$ can be used as a trial period τ_g. We could calculate $\mathcal{V}(\tau_g)$ by storing the pairs $(\langle n \rangle_{\tau_g}, x[n])$ for $n = 0, \ldots, N-1$, sorting them by abscissas $\langle n \rangle_{\tau_g}$, then using the resulting composite period $x_{\tau_g}[m]$ in Equation 3.1. All samples $x_{\tau_g}[m]$ would have distinct ordinates. For the trial period between the same two critical periods that surround the true period τ_w, the samples would be in the correct order, as well.

However, if we choose a rational value u/v $((u,v) = 1)$ for τ_g which also is not a critical period, it is possible to avoid storing and sorting the location/value pairs before calculating each $\mathcal{V}(\tau_g)$. In addition, the composite period samples will have integral locations. We can calculate $\mathcal{V}(\tau_g)$ by determining which samples succeed one another in the composite period and alternately accumulating successive absolute differences.

The samples $x[n]$ are to be sorted on $\langle n \rangle_{u/v}$ or equivalently, $\langle vn \rangle_u$, since temporal ordering is independent of time scale. There will thus be u samples in the composite period. Since u/v cannot be a critical period, $u \geq N$, and there will be $u - N$ missing samples in the composite period. If we were to use the original method of storing and sorting, we would provide u empty registers, then fill them with the N samples $x[n]$. The $u - N$ registers which would remain empty would be skipped in computing $\mathcal{V}(\tau_g)$ with Equation 3.1.

It is desirable to avoid providing the u registers for accumulating the absolute differences of successive composite period samples. We only need to determine which sample $x[m]$ would have been placed in register $r + 1$, given that sample $x[p]$ would have been placed in register r. We know that

$$\langle vp \rangle_u = \langle r \rangle_u$$

and

$$\langle vm \rangle_u = \langle r + 1 \rangle_u$$

Therefore,

$$\langle vp - 1 \rangle_u = \langle vm \rangle_u$$
$$\langle v(m - p) \rangle_u = 1$$

and

$$\langle m \rangle_u = \langle p + s \rangle_u \tag{3.20}$$

where s is the *multiplicative inverse* of v for the modulus u, i.e., any solution of

$$\langle vs \rangle_u = 1 \tag{3.21}$$

The complete set of solutions s is a residue class[5] of the modulus u, though the unique value less than u will be used. A method for solving Equation 3.21 will be presented near the end of this section.

To compute $\mathcal{V}(\tau_g)$, we store the first input sample, $x[0]$, then alternate between determining the next composite period sample with Equation 3.20, and accumulating the difference between it and the last sample stored. Whenever the index of the next sample is greater than the number available, a sample will be missing from the composite period. We simply skip this sample by determining the next sample and retaining the previously stored sample, since missing samples should not contribute to the variation. Computation of $\mathcal{V}(\tau_g)$ terminates when the next input sample is $x[0]$, i.e., the starting sample. We then will have alternated as above u times and accumulated N absolute differences.

Each composite period formed using $\tau_g = u/v$ will contain u uniformly-spaced samples. As mentioned earlier, $u - N$ samples will be missing. We therefore should choose τ_g whose numerators are as small as possible ($= N$, ideally). Recall that in searching for critical periods, we might find Farey fractions y/x between them. Since these values are all in lowest terms (i.e., $(x, y) = 1$), any of them is convenient for use as a trial period. Therefore, if more than one are found between a pair of critical periods, we should choose the one with the smallest numerator y. Doing so has the additional benefit of accelerating the variation computation without affecting the value obtained.

[5]See section 2.3.

If no Farey fractions (of the order given by Equation 3.9) exist between two particular critical periods a/b and c/d, we must find some other rational number u/v between them for τ_g. The *mediant* [7] of a/b and c/d, u/v, provides a convenient solution:

$$\begin{aligned} u &= a + c \\ v &= b + d \end{aligned} \tag{3.22}$$

Clearly,

$$\frac{a}{b} < \frac{u}{v} < \frac{c}{d}$$

In addition, u/v is sure to be in lowest terms. To prove this, suppose that u and v have a common factor g. Then

$$\begin{aligned} a + c &= ge \\ b + d &= gf \end{aligned}$$

where e and f are some integers. Now

$$\begin{aligned} c &= ge - a \\ b &= gf - d \end{aligned}$$

and

$$cb = g^2ef - g(af + ed) + ad$$

Utilizing Equation 3.15, we note that

$$g^2ef - g(af + ed) = 1$$

i.e., g is an integer factor of one. Therefore, g must equal one, and $(u, v) = 1$.

Once we find $\tau_{best} = u/v$, a value of τ_g for which $\mathcal{V}(\tau_g)$ is smallest, we can reconstruct the analog waveform by storing the samples $x[n]$ in the same order that they were used in computing $\mathcal{V}(\tau_{best})$. In particular, the multiplicative inverse of v, modulo u (Equation 3.21) is the increment s for the index n. As before, each successive index must be interpreted modulo u. Missing samples, indicated by $n \geq N$, must be blanked. Since there are u samples in the reconstructed period and the true period T_w is very nearly $\tau_{best}T_s$ (in units of real time), the sample spacing is $\tau_{best}T_s/u$, or T_s/v.

The remainder of this section will contain descriptions of procedures presented by Rader for computing a multiplicative inverse and initializing the Farey sequence used

to generate trial periods. Finally, flowcharts summarizing our implementation of the entire algorithm will be provided.

To solve Equation 3.21 for s, the multiplicative inverse of v, modulo u (assuming $(u, v) = 1$), we begin by expressing u/v as a *continued fraction* [7,10]:

$$\frac{u}{v} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 - \ddots + \cfrac{1}{a_\mu}}}} \tag{3.23}$$

The integers a_i are determined by the following equations:

$$\begin{aligned}
\frac{u}{v} &= a_0 + \frac{r_0}{v} & 0 \le r_0 < v \\
\frac{v}{r_0} &= a_1 + \frac{r_1}{r_0} & 0 \le r_1 < r_0 \\
\frac{r_0}{r_1} &= a_2 + \frac{r_2}{r_1} & 0 \le r_2 < r_1 \\
\frac{r_1}{r_2} &= a_3 + \frac{r_3}{r_2} & 0 \le r_3 < r_2 \\
&\vdots \\
\frac{r_{\mu-2}}{r_{\mu-1}} &= a_\mu
\end{aligned} \tag{3.24}$$

The continued fraction expansion of any rational number u/v has to terminate (i.e., $r_\mu = 0$) since each remainder r_i must be a non-negative integer smaller than its predecessor, r_{i-1}.

The expressions

$$\begin{aligned}
\frac{p_0}{q_0} &= \frac{a_0}{1} \\
\frac{p_1}{q_1} &= a_0 + \frac{1}{a_1} \\
\frac{p_2}{q_2} &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}} \\
&\vdots \\
\frac{p_\mu}{q_\mu} &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots + \cfrac{1}{a_\mu}}}} = \frac{u}{v}
\end{aligned} \tag{3.25}$$

where $(p_i, q_i) = 1$ for $i = 0, \ldots, \mu$, are called *convergents* to the continued fraction in Equation 3.23. It can be shown [7] that the even convergents p_{2i}/q_{2i} are all $\leq u/v$ and increase strictly with i, and the odd convergents p_{2i+1}/q_{2i+1} are all $\geq u/v$ and decrease strictly with i. Therefore, increasing values of i yield successively better approximations of u/v. The last convergent p_μ/q_μ is identically u/v.

For $n \geq 2$, the convergents can be generated iteratively [7,10] using

$$\begin{aligned}
p_n &= a_n p_{n-1} + p_{n-2} \\
q_n &= a_n q_{n-1} + q_{n-2}
\end{aligned} \tag{3.26}$$

In addition,

$$q_n p_{n-1} - p_n q_{n-1} = (-1)^n \tag{3.27}$$

Equations 3.26 and 3.27 can also be used for $n = 1$ if we define

$$\frac{p_{-1}}{q_{-1}} = \frac{1}{0} \tag{3.28}$$

We can now specify a procedure for solving Equation 3.21. Store the first two convergents: $p_{-1} = 1$, $q_{-1} = 0$, $p_0 = a_0$ $(= \lfloor u/v \rfloor)$, and $q_0 = 1$. Use Equation 3.24 to compute the integers a_n until a remainder r_n is zero. Also, as each a_n is calculated, compute the next pair (p_n, q_n) using Equation 3.26. If the zero remainder is found when

$n = \mu$, then $p_\mu = u$ and $q_\mu = v$. Using these values in Equation 3.27 and interpreting both sides modulo u, we see that

$$\langle v p_{\mu-1} \rangle_u = \langle (-1)^\mu \rangle_u \tag{3.29}$$

If μ is even, then $p_{\mu-1}$ is the solution to Equation 3.21:

$$s = p_{\mu-1} \tag{3.30}$$

If it is odd, then multiply both sides of Equation 3.29 by -1, and again interpret the results modulo u. This yields the solution

$$s = u - p_{\mu-1} \tag{3.31}$$

Several of the results presented above for determining a multiplicative inverse can also be used for initializing the Farey fraction generator needed to produce trial periods. We desire two consecutive Farey fractions of order L (Equation 3.9) spanning τ_{min}, the lower limit of the signal period search range:

$$\frac{u_1}{v_1} \leq \tau_{min} < \frac{u_2}{v_2} \qquad v_1, v_2 \leq L \tag{3.32}$$

We begin by noting the similarity between Equation 3.27 and Equation 3.15. Successive convergents p_{n-1}/q_{n-1} and p_n/q_n generated using Equation 3.26 are also adjacent[6] Farey fractions of some order L, where L satisfies Equations 3.12 and 3.16, i.e.,

$$\begin{aligned} q_{n-1} &\leq L \\ q_n &\leq L \end{aligned} \tag{3.33}$$

and

$$q_{n-1} + q_n > L \tag{3.34}$$

Equation 3.26 indicates that both the numerators and denominators of successive convergents increase with increasing n, though not necessarily by constant increments. Therefore, we can generate convergents to the continued fraction expansion of τ_{min},

[6] We use the term adjacent rather than consecutive since the two Farey fractions are in either ascending or descending order, depending on whether n is odd or even, respectively.

starting with a denominator of zero (q_{-1}), until a denominator $q_n \geq L$ is found. Rader has shown that the last two convergents p_{n-1}/q_{n-1} and p_n/q_n found in this manner provide the two desired Farey fractions, either directly or with a few additional, minor steps.

The only difficulty we may encounter is that a remainder of zero (see Equation 3.24) might be obtained before the termination condition above is met. If τ_{min} is irrational, this problem cannot occur. Unlike that of a rational number u/v, the continued fraction expansion of an irrational number is infinite. The integers a_i in Equation 3.23 (in which we replace u/v with τ_{min}) are unique, and are computed using the greatest integer function:

$$\begin{aligned} a_0 &= \lfloor \tau_{min} \rfloor \\ a_1 &= \left\lfloor \frac{1}{\tau_{min} - a_0} \right\rfloor \\ a_2 &= \left\lfloor \frac{1}{\frac{1}{\tau_{min} - a_0} - a_1} \right\rfloor \\ &\vdots \end{aligned} \tag{3.35}$$

In practice, rational numbers with many non-trivial digits must be used for τ_{min}. If a zero remainder is obtained before a value $q_n \geq L$ is found, we must adjust τ_{min} by some arbitrarily small amount ϵ, and restart.

If the last convergent q_n equals L, then by Equations 3.33 and 3.34, p_{n-1}/q_{n-1} and p_n/q_n are the desired Farey fractions. If $q_n > L$, then p_n/q_n cannot be a Farey fraction of order L. However, it will be shown that p_{n-1}/q_{n-1} is still one of the two we seek, and that the other (p'/q') is given by

$$a' = \left\lfloor \frac{L - q_{n-2}}{q_{n-1}} \right\rfloor \tag{3.36}$$

and

$$\begin{aligned} p' &= a' p_{n-1} + p_{n-2} \\ q' &= a' q_{n-1} + q_{n-2} \end{aligned} \tag{3.37}$$

The denominator q' must be $\leq L$ since

$$a' \leq \frac{L - q_{n-2}}{q_{n-1}}$$

We know that the denominator of p_{n-1}/q_{n-1} is $\leq L$. (If this was not the case, convergent generation would have had to terminate earlier.) Since

$$q_{n-1}p_{n-2} - p_{n-1}q_{n-2} = (-1)^{n-1}$$

we can easily verify that

$$q'p_{n-1} - p'q_{n-1} = (-1)^{n}$$

which meets the requirement imposed by Equation 3.15.

Finally, we see from Equations 3.36 and 3.37 that Equation 3.16 is also satisfied since

$$q' + q_{n-1} = q_{n-1}(1 + \left\lfloor \frac{L - q_{n-2}}{q_{n-1}} \right\rfloor) + q_{n-2}$$

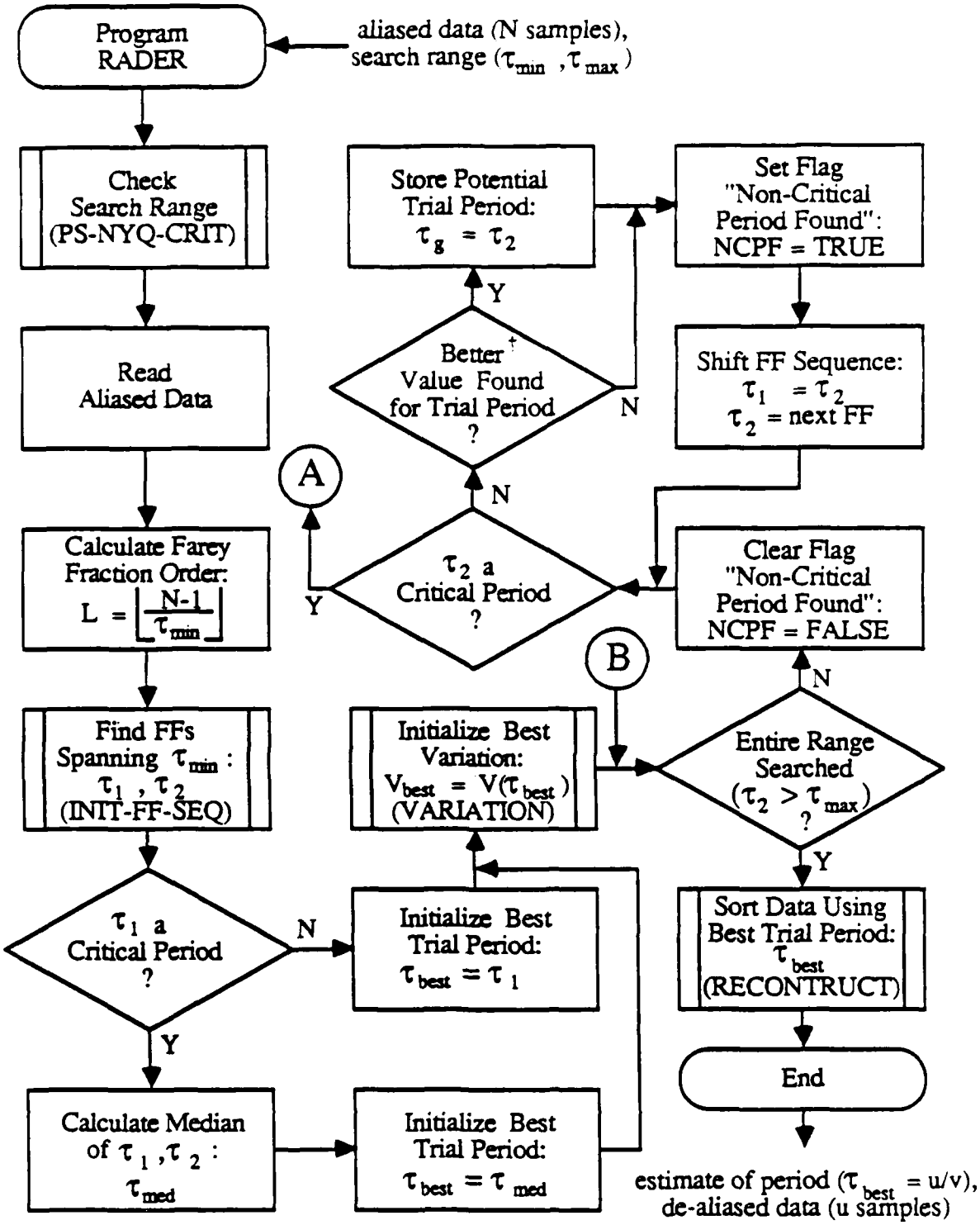
and $1 + \lfloor r \rfloor > r$ for any r.

Figures 3.6 through 3.11 contain flowcharts summarizing our implementation of the procedures reviewed in this section. They comprise a main program (RADER) and five subroutines, four of which are called directly from the main program. Subroutine calls are indicated by boxes with two additional vertical lines.

3.3 A Modification of the Algorithm

The success of the Rader algorithm in recovering a given aliased signal depends largely on N, the number of samples used. This can be inferred from the probabilistic arguments supporting the minimum variation criterion which were given in Section 3.1. N indirectly determines the density of the search for τ_w along the τ-axis. The density increases directly (though non-uniformly) with increasing Farey fraction order L. Since the search range lower limit τ_{min} for a given signal must be known beforehand, L is determined by (and approximately proportional to) N, as evident from Equation 3.9.

If too few samples are used, then the algorithm will fail. Specifically, for a given τ_{min}, there is an (approximate) minimum number of samples M yielding a search density insuring that the estimated period τ_{best} and the true period τ_w both lie between the same two critical periods in the corresponding Farey series. However, M is impossible to quantify, and we must proceed under the assumption that enough samples will be



Notes:

† Either τ_2 is first candidate for next trial period τ_g (i.e., NCPF = FALSE), or numerator of τ_2 < numerator of previously stored candidate (if NCPF = TRUE). Using new value will result in fewer missing composite period samples.

Figure 3.6: Program RADER.

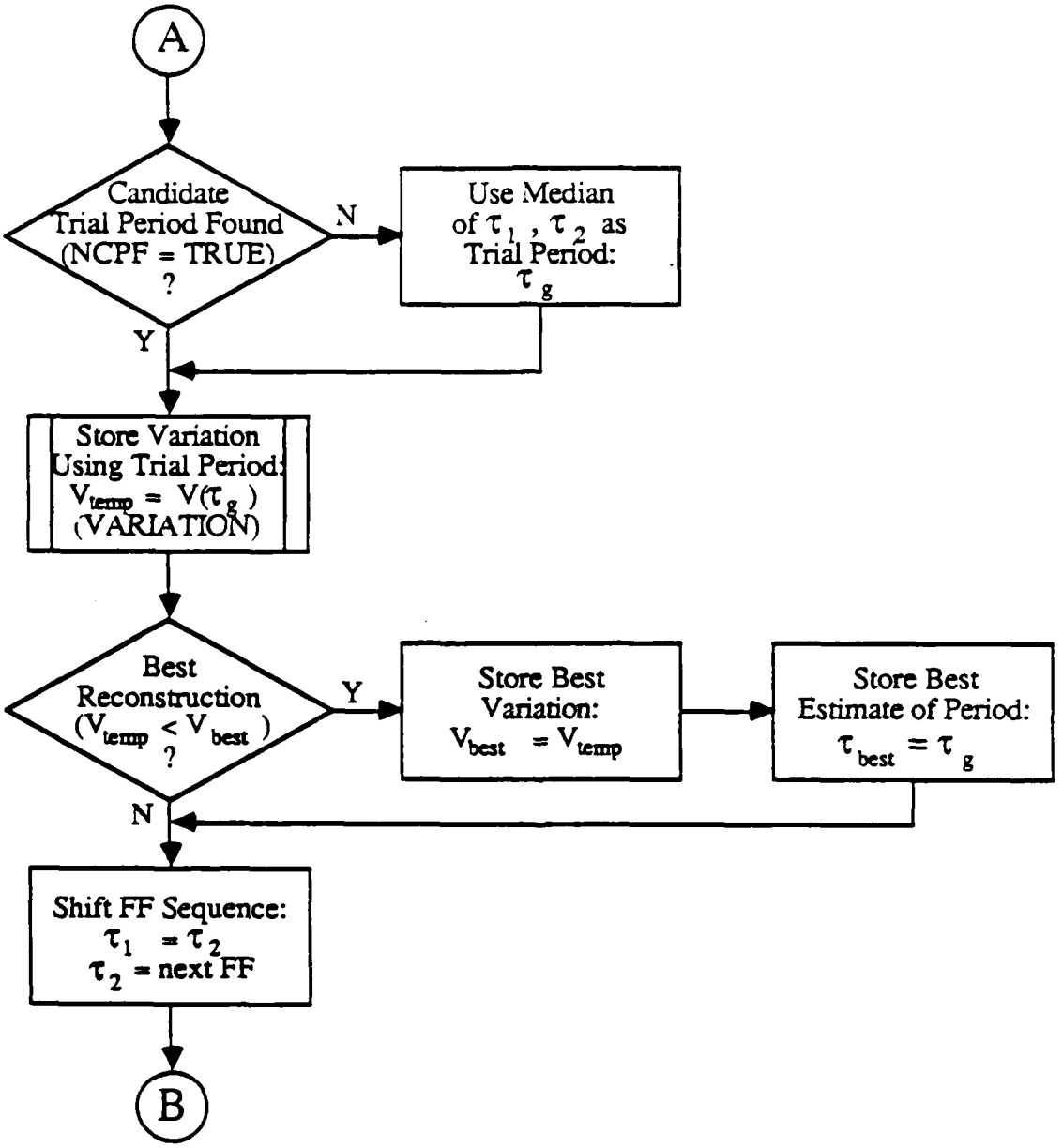
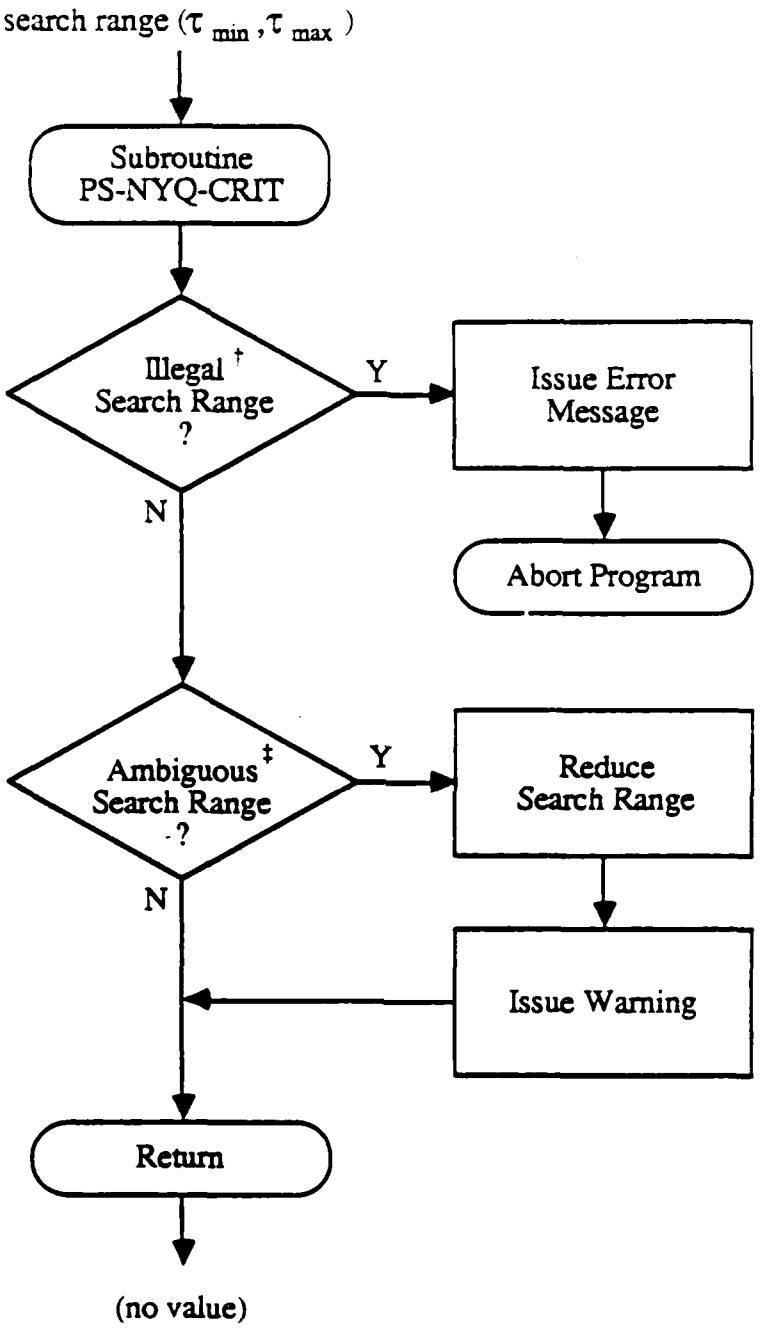


Figure 3.6: *continued*



Notes:
† Negative τ_{min} or τ_{max}, or $\tau_{min} > \tau_{max}$.
‡ One which does not satisfy the pseudo-Nyquist criterion.

Figure 3.7: Subroutine PS-NYQ-CRIT.

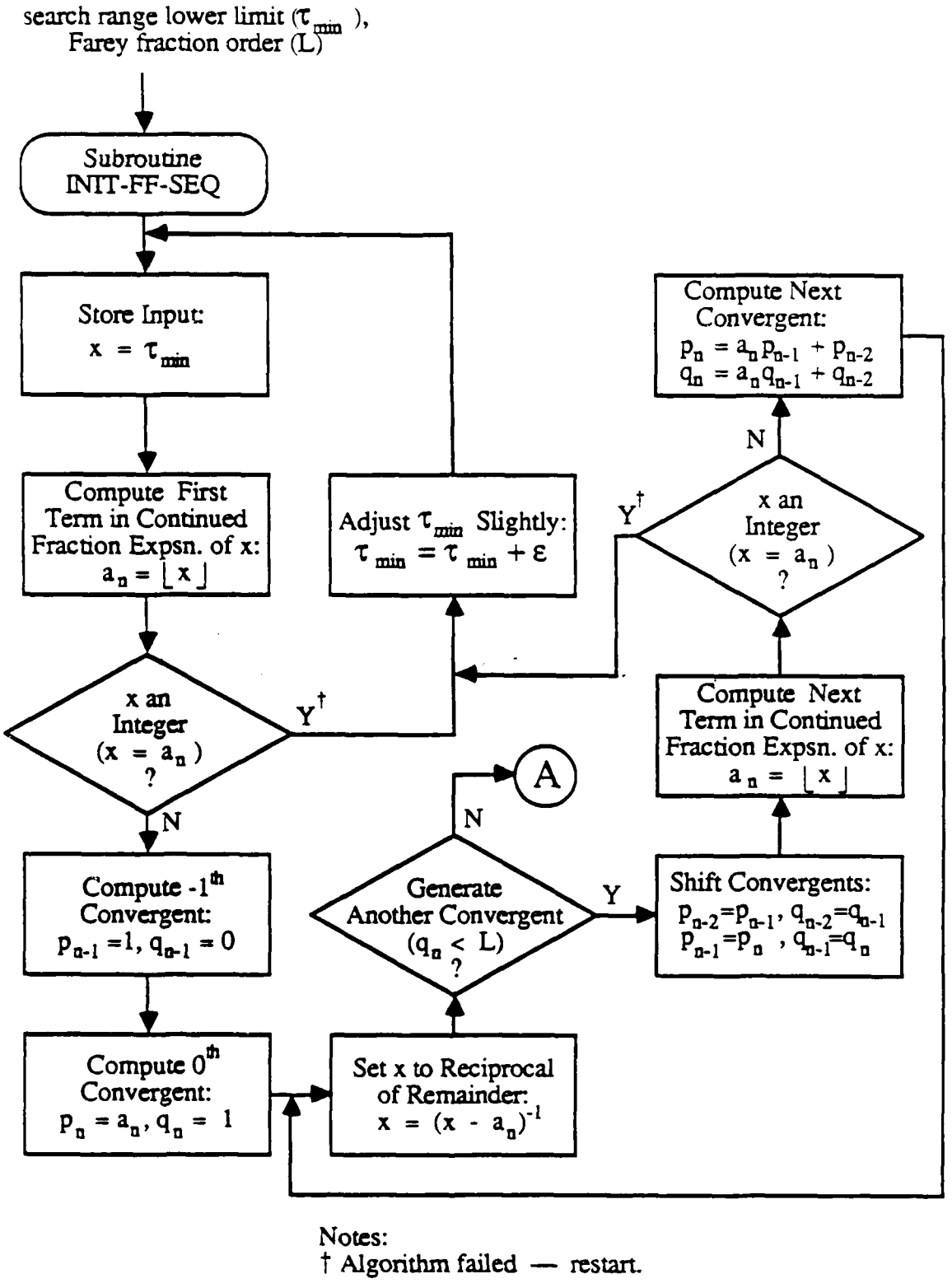
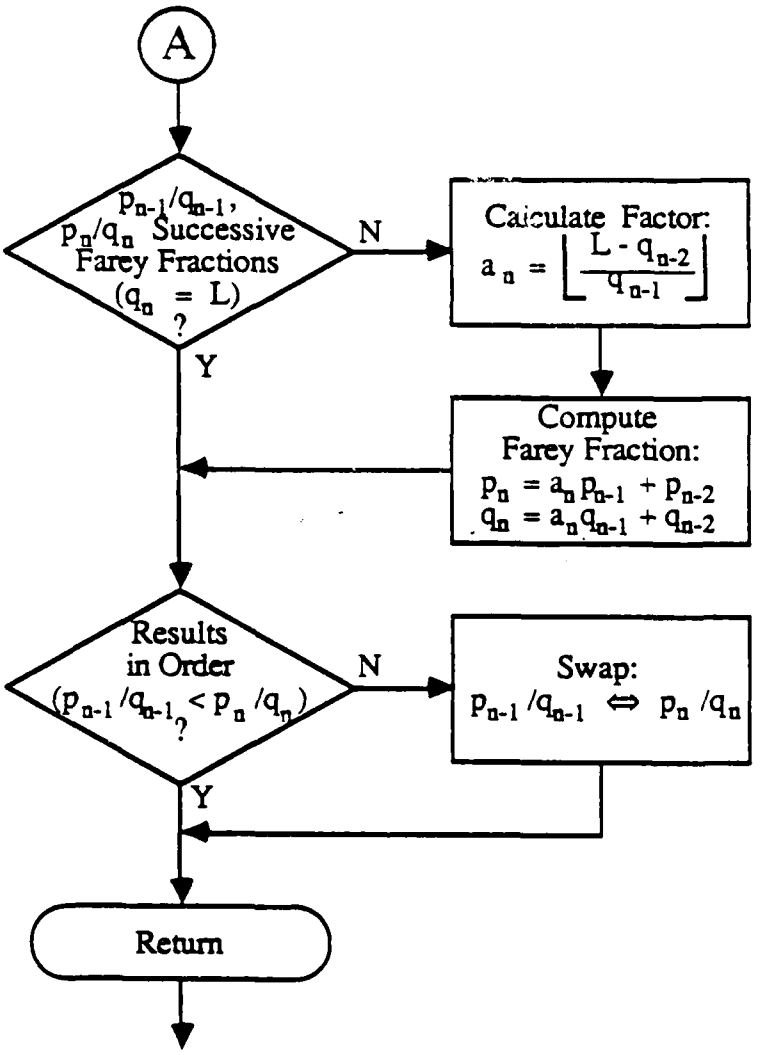


Figure 3.8: Subroutine INIT-FF-SEQ.



Farey fractions of order L
(p_{n-1}/q_{n-1}, p_n/q_n) spanning τ_{min} .

Figure 3.8: *continued*

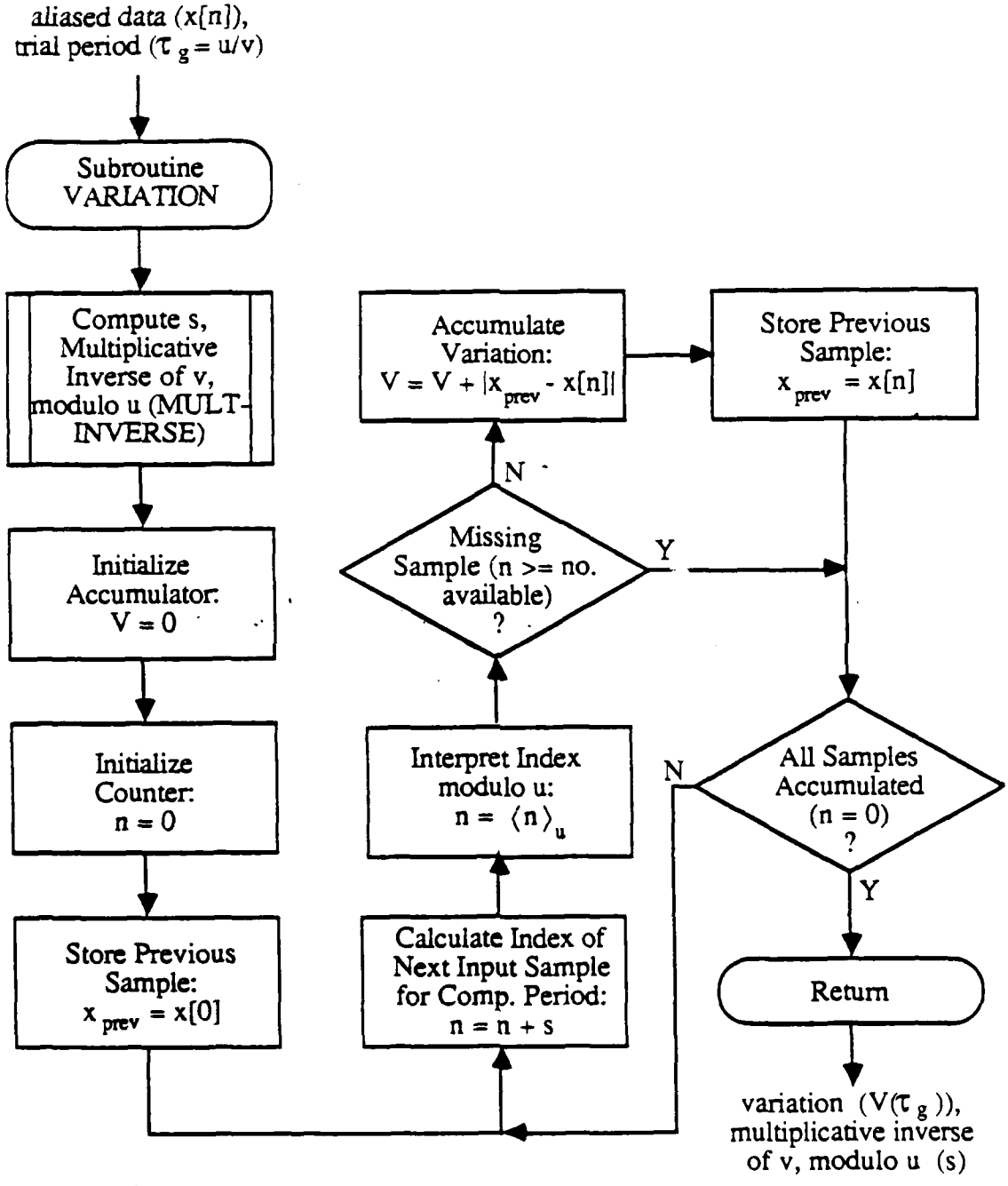
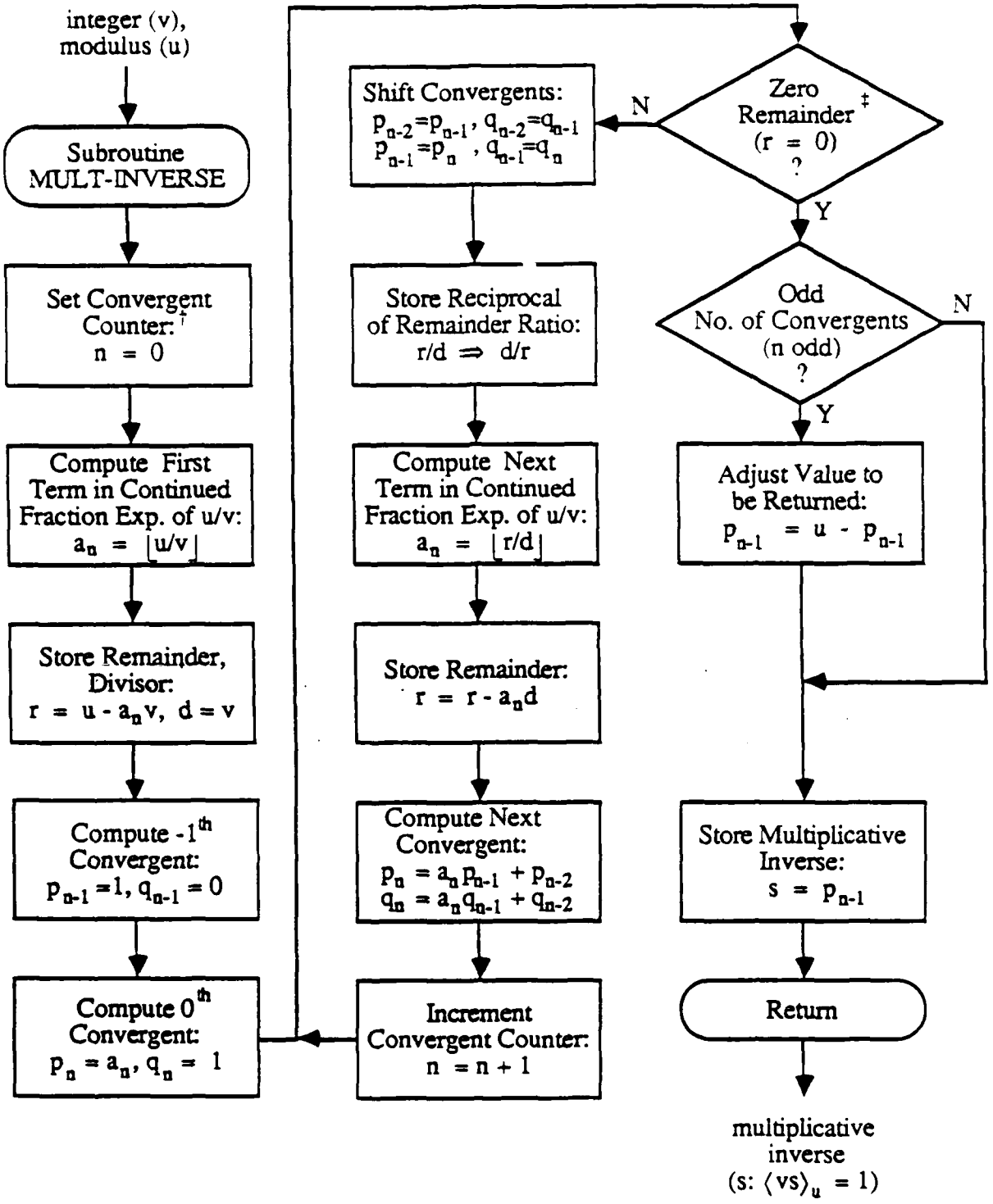


Figure 3.9: Subroutine VARIATION.



Notes:
† Equals number of terms in continued fraction expansion of u/v.
‡ Never true on first iteration (greatest common denominator (u, v) cannot = 1).

Figure 3.10: Subroutine MULT-INVERSE.

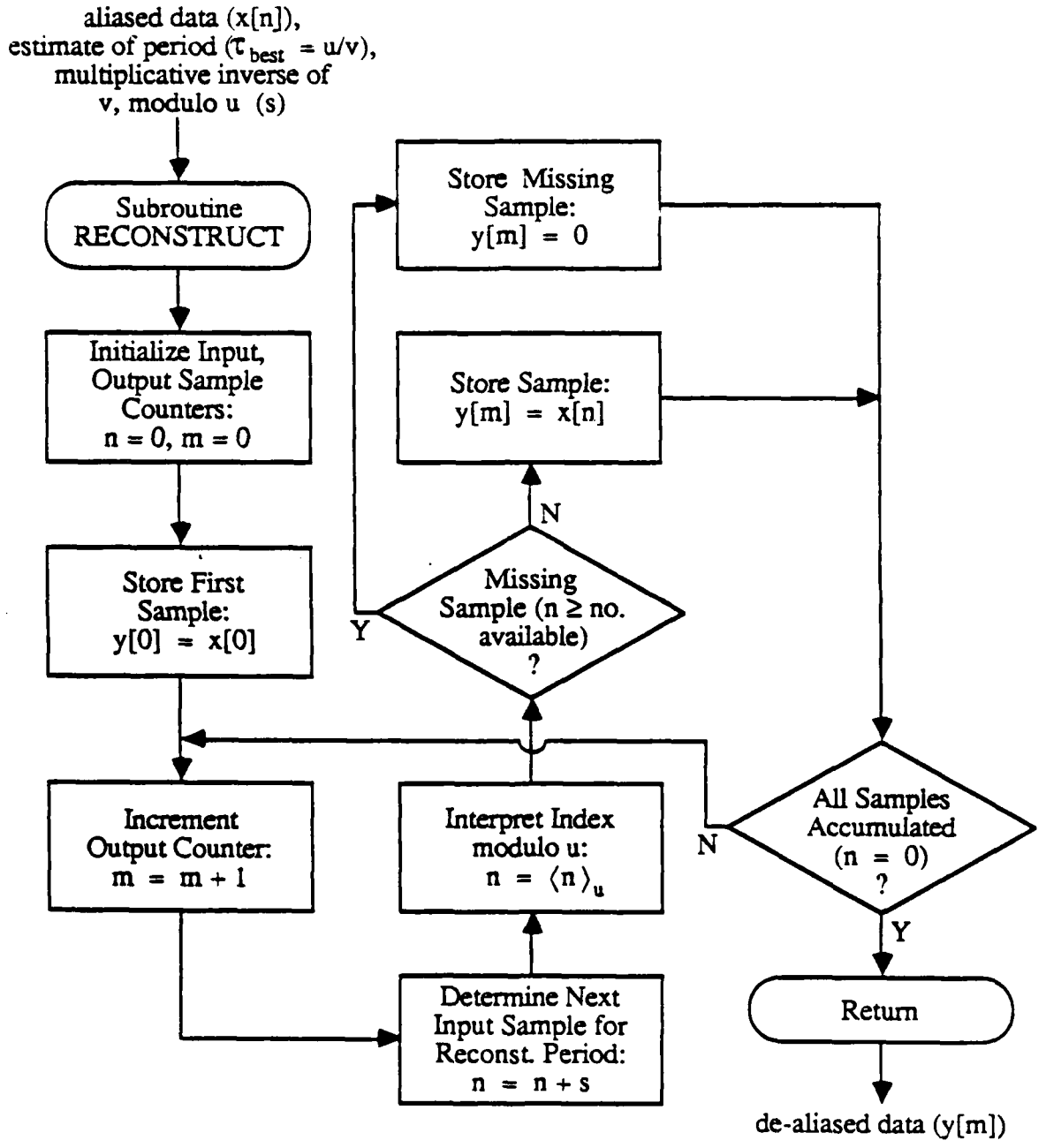


Figure 3.11: Subroutine RECONSTRUCT.

used. The minimum variation criterion presupposes the latter. If N is much greater[7] than the number of significant harmonics in the waveform, then the algorithm should not fail, and τ_{best} should approximately equal τ_w. We shall continue to assume that N is always sufficiently large.

We now consider the effect of using a quantity of samples substantially exceeding the unknown minimum M. This should increase the accuracy of the estimate of τ_w and yield better waveform reconstruction. However, the computational expense incurred is generally very high since the order of the search time[8] is much greater than $O\{N\}$. The search time consists primarily of the time required for computing composite period variations using the trial periods. The number of trial periods is equal to the number of critical period Farey fractions in the search range.

Suppose that the number of samples available is substantially greater than M. Assuming M can be estimated *very* roughly (perhaps from the number of harmonics postulated), it seems probable that the minimum acceptable search density can be used in an initial search to substantially reduce the uncertainty range of τ_w. This range, over which the corresponding variation function would be constant (and minimized, as well), would be delimited by the two critical periods spanning the returned estimate τ_{best}. A reduced quantity of samples (N_R), slightly greater than M, would be used for this coarse search. Successively finer searches, requiring increasing N_R, could then be used to reduce the uncertainty and refine the estimate of τ_w further. This iterative procedure would be terminated once either all samples had been used, the estimate was sufficiently accurate, or the additional processing time became prohibitive.

In order to insure that the search density increases monotonically between iterations, we must increment the Farey fraction order, rather than directly incrementing the number of samples. We choose an initial Farey fraction order L_o, and a Farey fraction order increment ΔL. Equation 3.9 can be rewritten to show the relationship between

[7] By perhaps an order of magnitude or more.

[8] This quantity is difficult to enumerate due to the irregular distribution of Farey fractions.

the Farey fraction order and the reduced number of samples used on an iteration:

$$L = \left\lfloor \frac{N_R - 1}{\tau_{Rmin}} \right\rfloor \tag{3.38}$$

τ_{Rmin} is the lower limit of the reduced search range on a given iteration, and will be specified below. Directly increasing N_R does not insure monotonically-increasing L since τ_{Rmin} also increases monotonically.

Given τ_{Rmin}, we compute the reduced number of samples necessary to yield the desired search density (approximately):

$$N_R = L\tau_{Rmin} + 1 \tag{3.39}$$

If this value exceeds N, the number available, then all samples are used (i.e., $N_R = N$) for one final iteration. Once N_R is determined, the Farey fraction order must be recomputed using Equation 3.38, due to the greatest integer function involved, and also since N_R may be limited, as above.

The search range for the initial iteration is the original, full search range: $\tau_{Rmin} = \tau_{min}$ and $\tau_{Rmax} = \tau_{max}$. On each iteration, the two critical periods spanning the estimate τ_{best} are retained to be used as the reduced search range $[\tau_{Rmin}, \tau_{Rmax}]$ on the next iteration. Since a Farey series of order L includes all members of an order-$(L-1)$ Farey series, and N_R must increase between iterations (since τ_{Rmin} can never decrease), we know that these two critical periods will be among the critical periods in the next iteration. (Recall that critical periods are Farey fractions whose numerators are less than N_R.)

On the first search, initialization of the trial period generator consists of determining the two order-L Farey fractions (τ_1 and τ_2) spanning the lower search range limit τ_{Rmin} (in this case, $= \tau_{min}$), in the same manner as in the original algorithm (see Figure 3.8). However, the initialization algorithm typically requires some adjustment when given a rational value τ_{Rmin}. Thus on successive iterations, Farey fractions spanning the *original* lower limit τ_{min} (always $\leq \tau_{Rmin}$), rather than τ_{Rmin}, should be found. A negligible amount of computation is then necessary to generate successive Farey fractions using Equations 3.13 and 3.14 until a critical period greater than[9] or equal to τ_{Rmin} is found.

This value would correspond to τ_2. τ_1 would be the preceding (non-critical period) Farey fraction. The Rader algorithm would then proceed as described in the previous section.

Figures 3.12 through 3.14 contain flowcharts summarizing our implementation of our modifications to the Rader algorithm. Together, the main program FAST-SCAN and subroutine RADER-SRCH replace the main program RADER previously shown in Figure 3.6. In addition, another module (subroutine RAISE-INIT) has been added to implement the modification in trial period generator initialization. All other subroutines in Section 3.2 remain unchanged.

3.4 Examples

The Rader algorithm, as well as all other algorithms presented in this thesis, was implemented in the C programming language on a VAX-11/750 minicomputer[10] under the UNIX 4.2 BSD operating system.[11] Pairs of integers were used to represent rational quantities such as trial periods. Double precision floating point numbers were used for most other quantities (e.g., input and output samples, and variations).

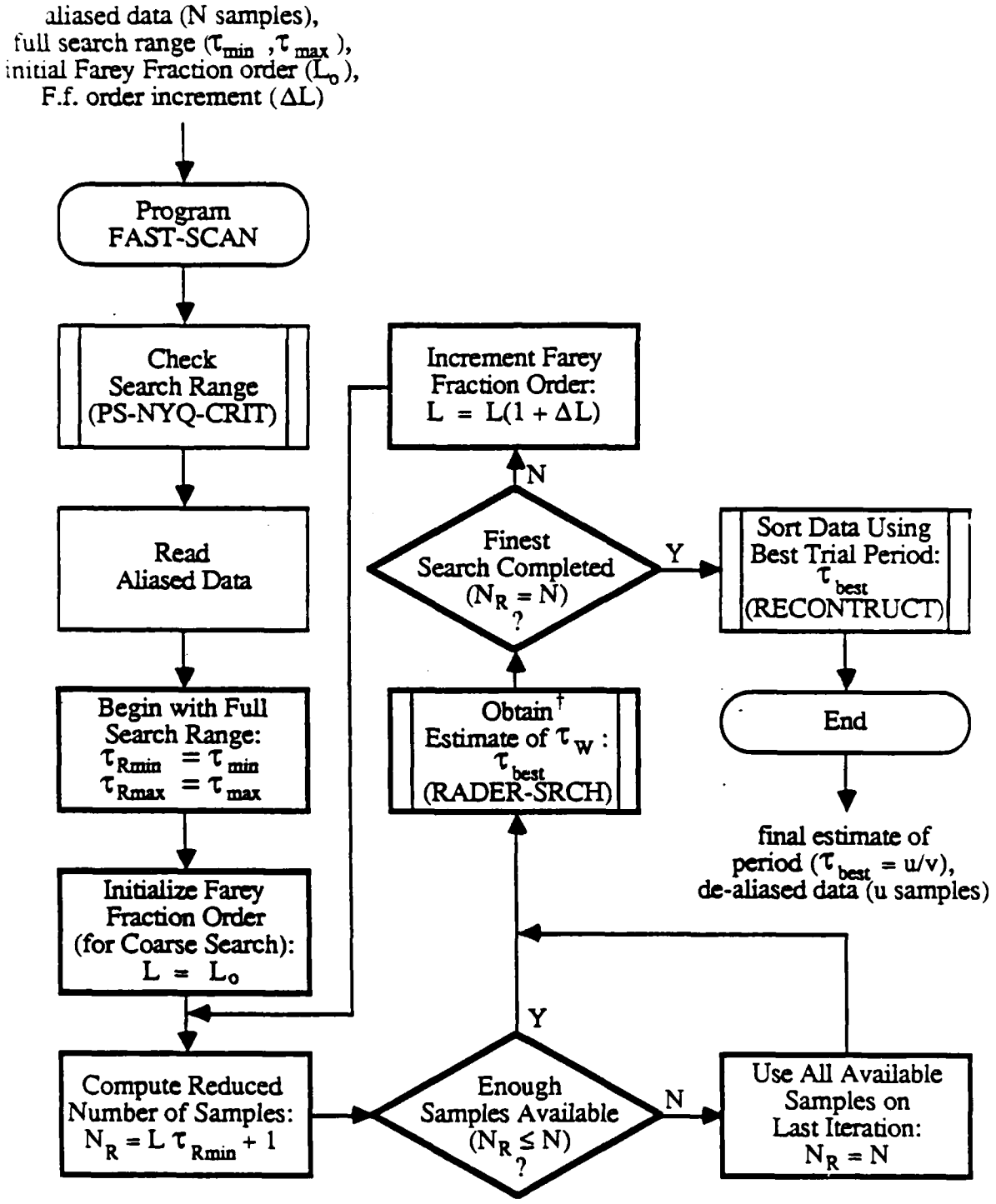
Figures 3.15 through 3.17 compare several recovered signals with their aliased counterparts. Oversampled signals are also shown for comparison, though their normalized time scales differ from those of the aliased and recovered versions, as noted in the captions. Figure 3.15 shows the simple case of a single sinewave originally having normalized frequency $\phi_w = .734531$, aliased to $\phi = .265469$ (with a phase shift of π). The frequency determined using the algorithm was .734375, in error by only .02%. Approximately 8 seconds were required to process 50 samples. The missing composite period samples are clearly evident in the plot.

Figure 3.16 contains a more harmonically rich, synthetic waveform comprising ten equal-amplitude sinusoids superimposed on a small d.c. offset. While the original signal

[9]The "greater than" applies only to the first iteration since τ_{Rmin} must be a critical period in all subsequent iterations.

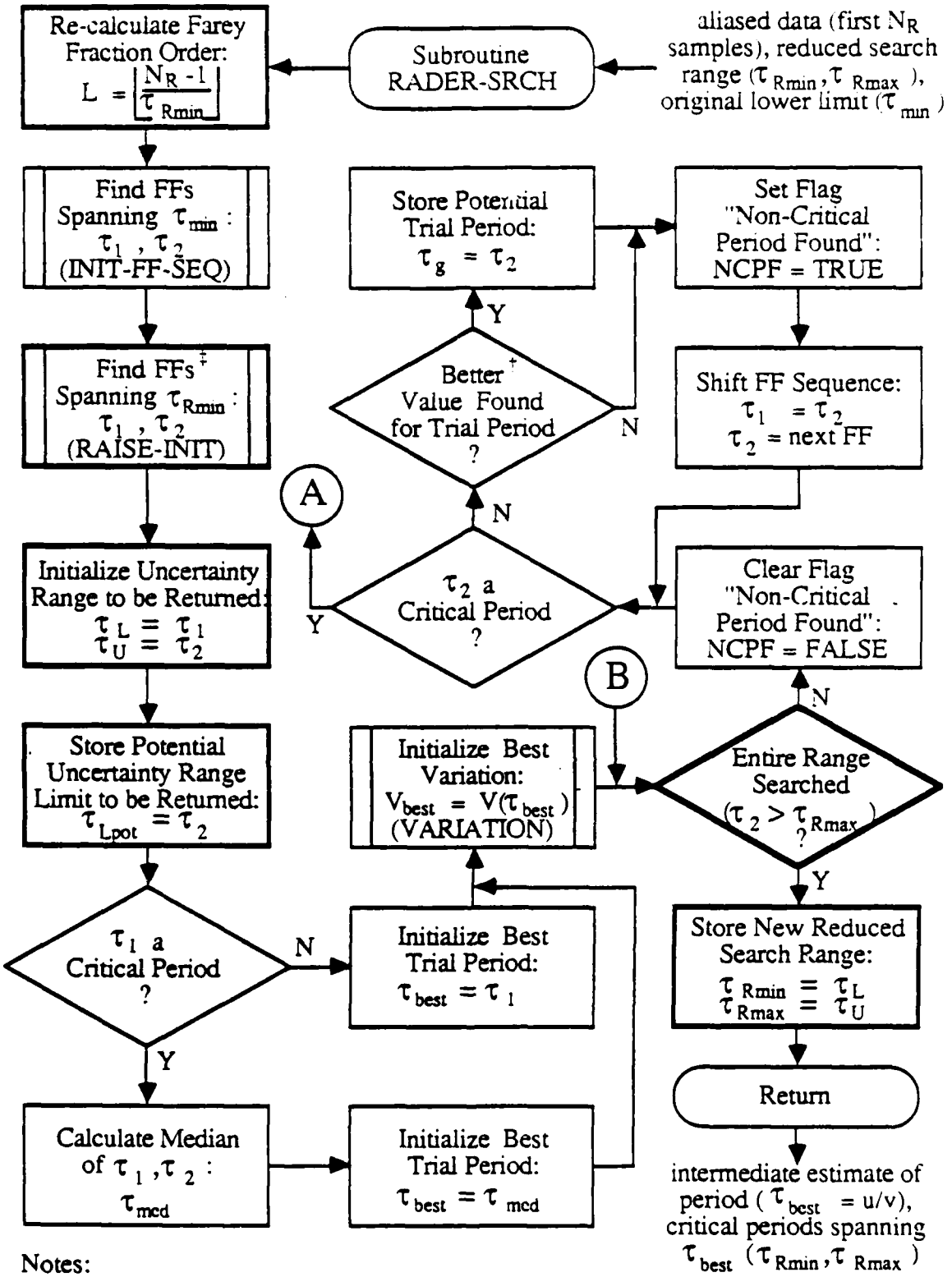
[10]VAX is a trademark of the Digital Equipment Corporation.

[11]UNIX is a trademark of AT&T Bell Laboratories.



Notes:
† Also determine new, narrower search range for next iteration.

Figure 3.12: Program FAST-SCAN. Bold blocks indicate additions to or modifications of Figure 3.6. (Also see Figure 3.13.)



Notes:

† Either τ_2 is first candidate for next trial period τ_g (i.e., NCPF = FALSE), or numerator of τ_2 < numerator of previously stored candidate (if NCPF = TRUE). Using new value will result in fewer missing composite period samples.

‡ τ_2 returned will be smallest critical period $\geq \tau_{Rmin}$.

Figure 3.13: Subroutine RADER-SRCH. Bold blocks indicate additions to or modifications of Figure 3.6. (Also see Figure 3.12.)

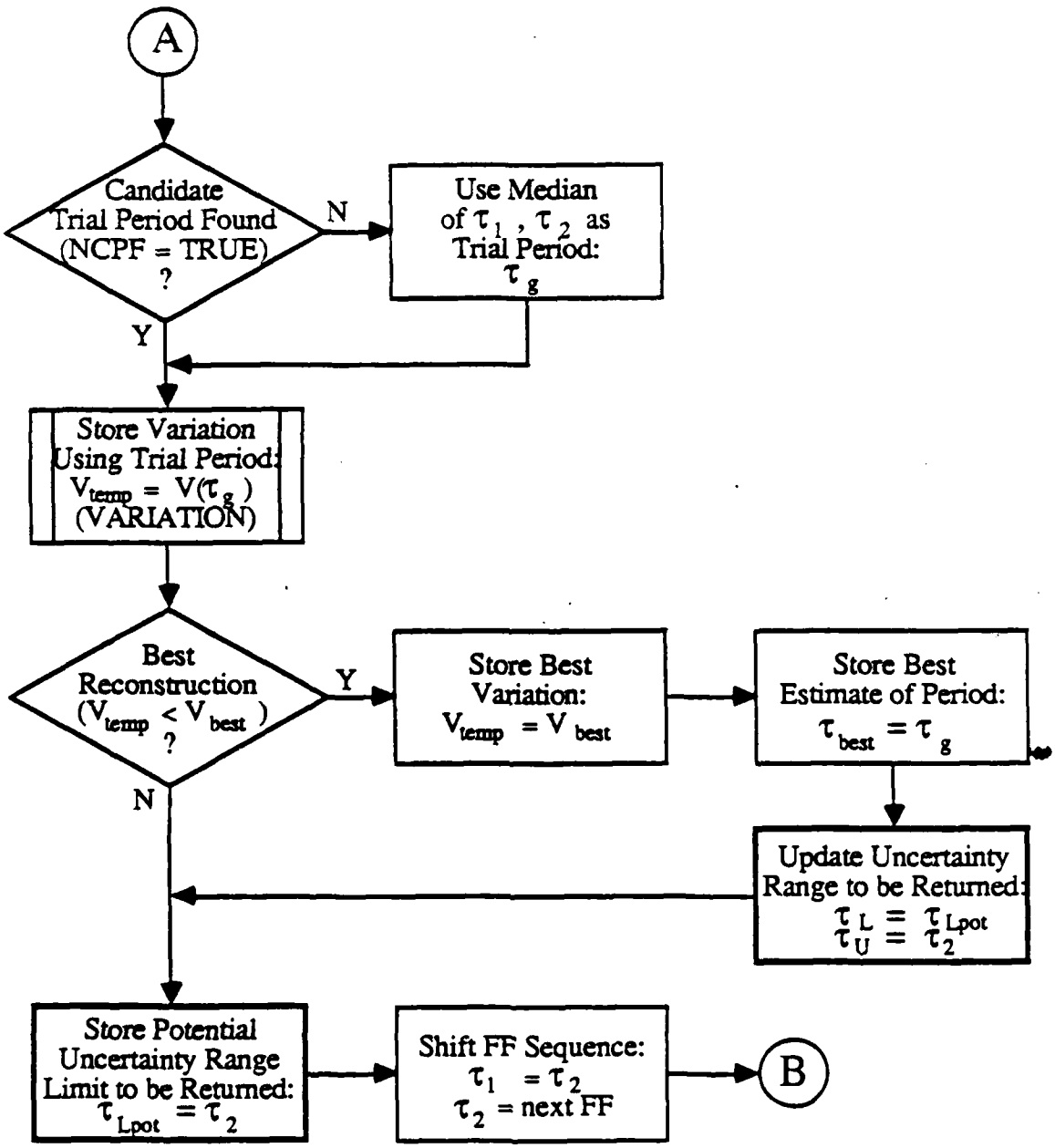
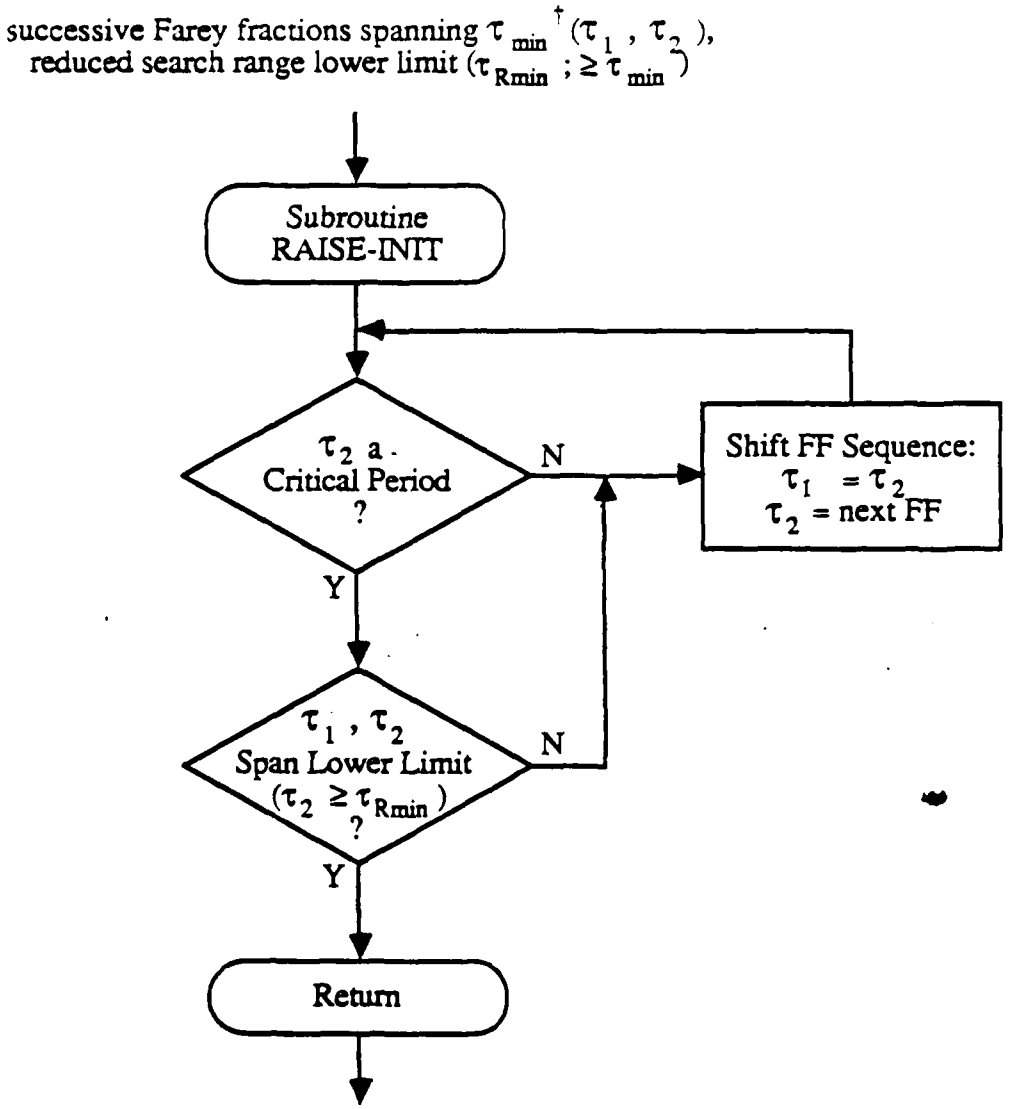


Figure 3.13: *continued*



Notes:
† Full search range lower limit.

Figure 3.14: Subroutine RAISE-INIT.

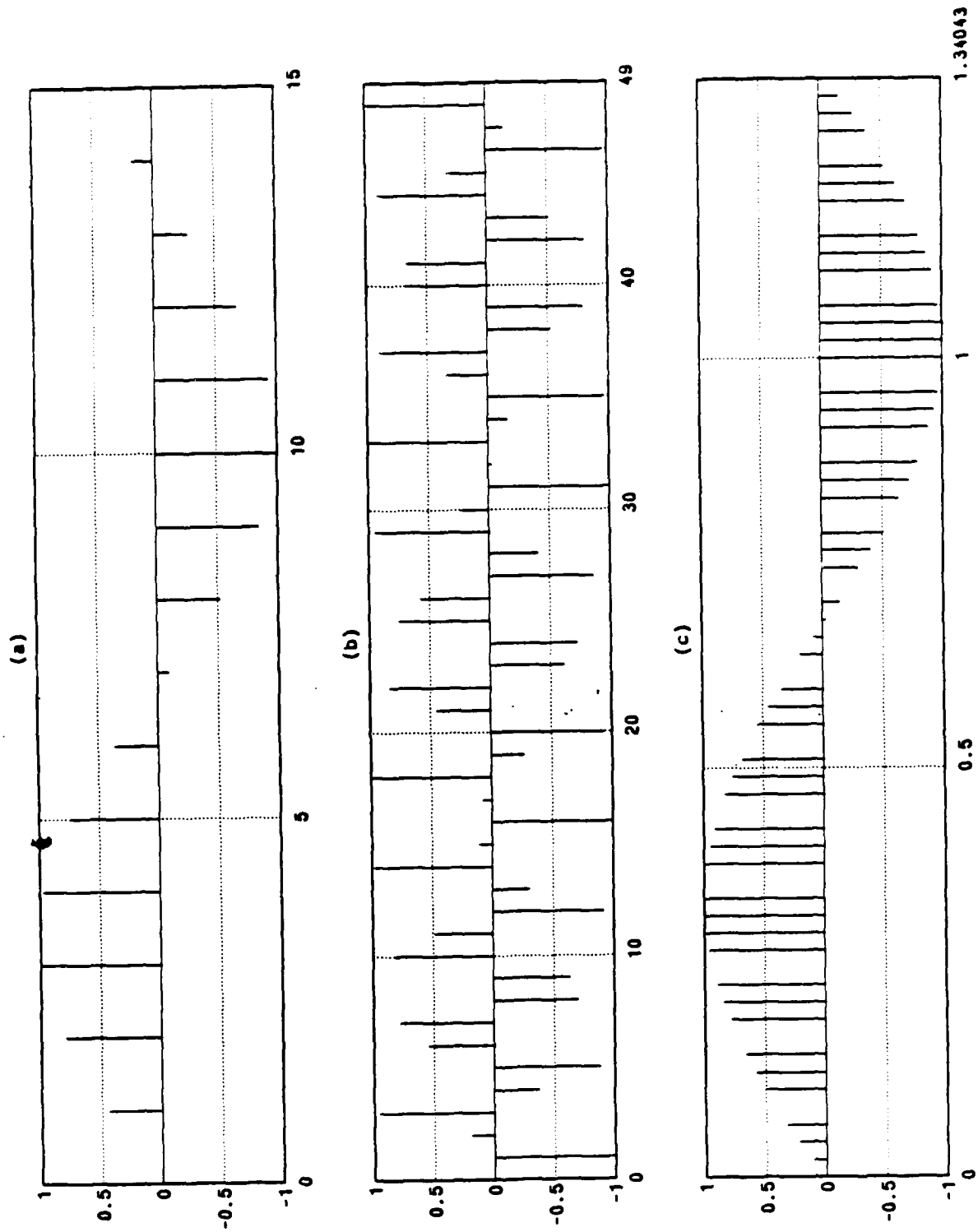


Figure 3.15: Aliased sinewave recovered using Rader algorithm. (a) Oversampled signal. (b) Aliased signal ((a) downsampled by 10). (c) Recovered signal (same normalized time scale as (b)).

is completely obscured by aliasing, as shown in Figure 3.16(b), it is readily apparent in Figure 3.16(c). 200 samples were used, requiring 146 seconds. Identical reconstructions were obtained using our FAST-SCAN version of the Rader algorithm with an initial Farey fraction order of 100 and an order increment of 10% per iteration. In this case, only 11 seconds were required.

The execution time is almost completely independent of the harmonic content of the waveform. However, we see from comparing the two previous examples of the unmodified Rader algorithm just how strongly the number of samples affects it. Experiments with 500 or more samples have confirmed this. Execution time will be discussed in greater detail in Chapter 5.

Figure 3.17 contains waveforms originating from 60 Hz line interference sampled at 10796.123 Hz. (This rate was chosen to meet the pseudo-Nyquist criterion irrationality requirement.) Identical reconstructions were obtained using the Rader (unmodified) and FAST-SCAN algorithms, the latter with an initial Farey fraction order of 250 and an order increment of 10%. 250 samples were used, requiring 551 and 15 seconds, respectively.

We conclude this chapter with a brief look at typical Rader algorithm variation functions $\mathcal{V}_N(\tau_g)$, where N is the number of samples used. Figure 3.18 shows three variation functions, all obtained from the aliased test signal of 10 sines which was used above. They differ only in the number of samples (N) used in each case. The most striking feature of the illustration is that the Rader algorithm converges to the true signal period τ_w as N is increased (though there is a slight undershoot for $N = 50$). In addition, $\mathcal{V}_N(\tau_g)$ continues to increase at all points $\tau_g \neq \tau_w$ as N is increased, while it quickly reaches a limiting value at $\tau_g = \tau_w$.

The widths of the uncertainty ranges, i.e., the intervals over which the variations are constant and minimized, decrease for increasing N. However, it is clear from the plot that $N = 25$ would be insufficient as the number of samples to use in the initial iteration of the FAST-SCAN algorithm since the resulting initial uncertainty range would not completely contain the proper uncertainty ranges for subsequent iterations. Nonetheless, FAST-SCAN would successfully determine τ_w if an initial iteration sample

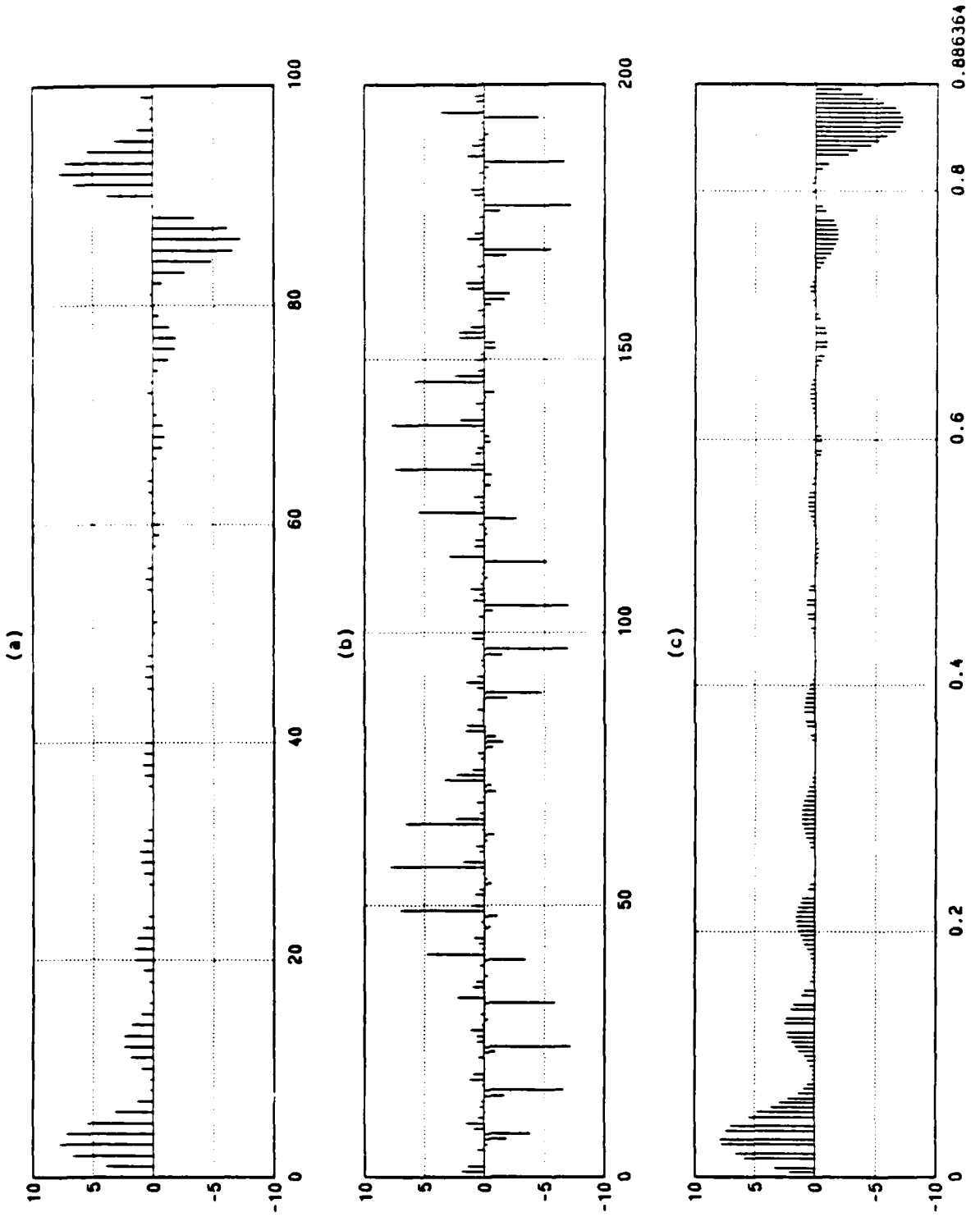


Figure 3.16: Aliased synthetic signal (ten sines with d.c. offset) recovered using Rader algorithm. (a) Oversampled signal. (b) Aliased signal ((a) downsampled by 100). (c) Recovered signal (same normalized time scale as (b)).

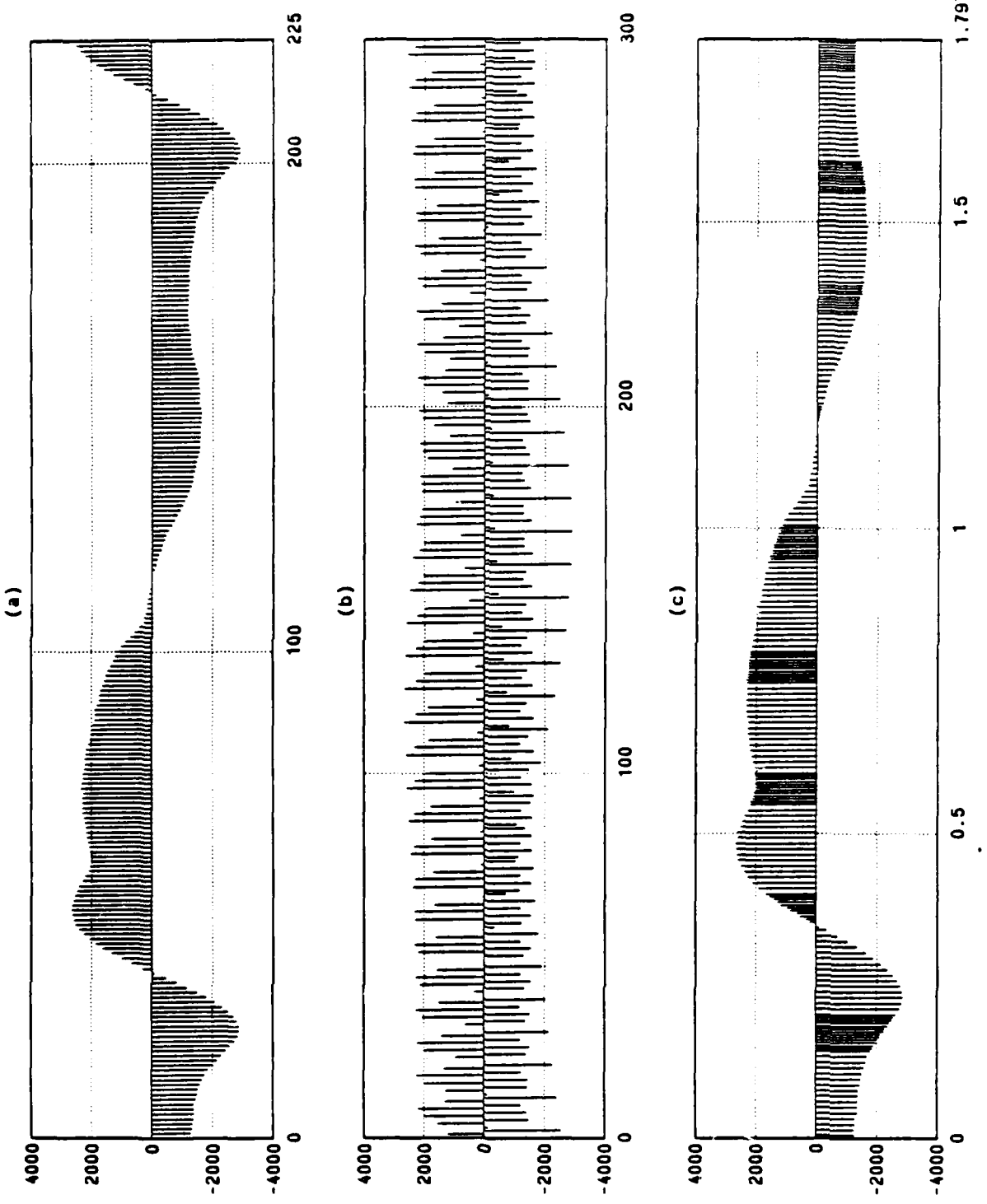


Figure 3.17: Aliased 60 Hz line interference signal recovered using Rader algorithm. (a) Oversampled signal. (b) Aliased signal ((a) downsampled by 100). (c) Recovered signal (same normalized time scale as (b)).

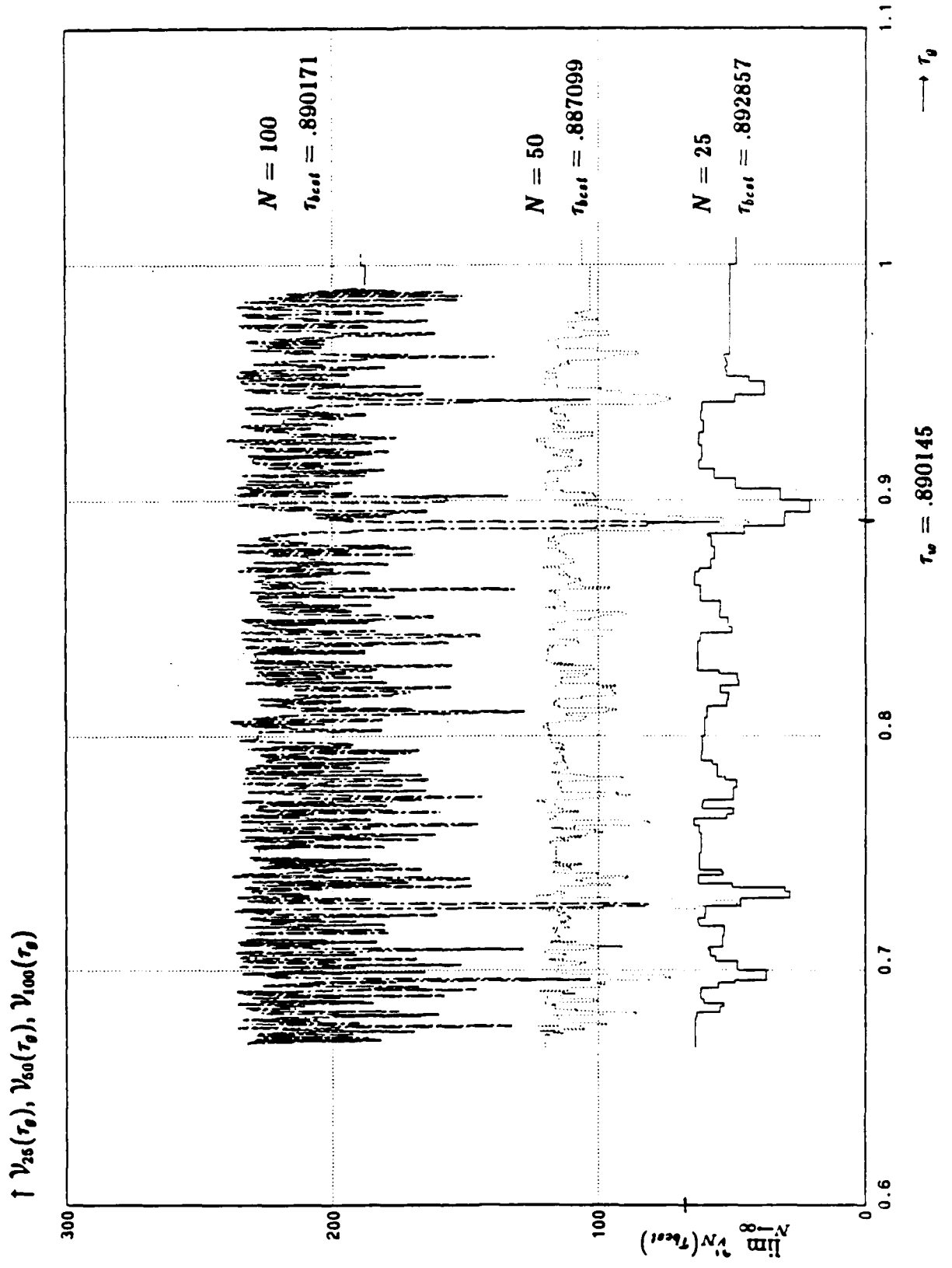


Figure 3.18: Convergence of variation function for increasing quantities of samples. Test signal from Figure 3.16.

count N_R of 50 or more was used.

Chapter 4

SPEC-PEAKS — A Frequency Domain Alternative to the Rader Algorithm

Like the Rader algorithm, the new algorithm to be presented in this chapter consists of estimating the normalized period[1] τ_w of an analog signal from aliased data, then reconstructing the signal by using the estimate in an appropriate data sorting routine. As before, a finite set of guesses is chosen, and some criterion is used to select the best one. However, the guesses will now be determined in the frequency domain. Specifically, we will use guesses of the fundamental frequency ϕ_w. Since they are analogous to the trial periods in the Rader algorithm, they will be called *trial frequencies*. Reconstruction will consist of sorting spectral samples corresponding to aliased harmonics using the best estimate of ϕ_w among the trial frequencies, then inverse transforming the results. Because the trial frequencies will be obtained by peak-picking the spectrum obtained from the discrete Fourier transform of the aliased data, we will refer to the proposed method as the spectral peaks algorithm, or simply "SPEC-PEAKS".

Development of the SPEC-PEAKS algorithm is closely related to the theoretical work in Chapter 2. Since we will often be able to draw upon this earlier work directly,

[1] Or equivalently, the normalized frequency, ϕ_w.

it will be appropriate to condense discussion of the general approach and a detailed description of the algorithm into a single section, Section 4.1. Section 4.2 will discuss a modification of the SPEC-PEAKS algorithm which, although not mandatory for signal recovery, will yield lower reconstruction error in certain cases. The closing section of the chapter will present examples of waveforms reconstructed using SPEC-PEAKS (with and without the modification), along with the corresponding oversampled and aliased signals.

4.1 General Approach and Detailed Description of the Algorithm

Whether or not the normalized signal fundamental frequency ϕ_w is known, the first step towards signal reconstruction comprises a discrete Fourier transform $X[k]$ of the aliased data $x[n]$. Our ultimate objective is to estimate the relative complex amplitudes of all significant harmonics in the original signal directly from the DFT samples. The aliased harmonics can then be sorted into a *composite spectrum*[2] whose inverse discrete Fourier transform provides one period of the original signal.

We begin with the simple case where ϕ_w is known precisely. The relative harmonic amplitudes are approximately equal to the DFT samples $X[k]$ which are nearest positive and negative multiples of the fundamental frequency. In order to yield the appropriate DFT indices, each multiple is first interpreted modulo the normalized sampling rate ϕ_s, which is unity by definition, and then scaled with the DFT length R:

$$k_i = \lfloor R\langle i\phi_w\rangle_1 \rceil \qquad i = 0, \pm 1, \pm 2, \ldots \tag{4.1}$$

The new notation $\lfloor\ \rceil$ has been introduced to represent the nearest integer or "rounding" function.

Since only a finite number of time samples are used to compute the DFT, the observed harmonics have measurable amplitudes. Observed harmonics never have zero width; therefore, we can expect that large quantities of samples will be needed to avoid

[2]Here again we choose notation which is consistent with that used by Rader.

destructive aliasing. If an insufficient number are used, adjacent aliased harmonics in the observed spectrum may overlap, even when the signal and sampling rates are not integrally related. Assuming overlap does not occur, the relative harmonic amplitudes found will be correct (to an arbitrary degree of accuracy, through zero-padding prior to the DFT), and we can proceed to sort the aliased harmonics in frequency using these values.

Harmonics from negative multiples of ϕ_w (Equation 4.1) are placed in the upper portion of a composite spectrum[3] $X_{\phi_w}[j]$, and the remainder in the lower portion. For simplicity, we hereafter will assume that there are equal numbers of significant harmonics in the original signal at positive and negative analog frequencies (excluding d.c.), unless the signal is analytic. If a total of P harmonics including the one at d.c. are localized using Equation 4.1, then they are sorted as follows:

$$\begin{aligned}
X[k_0] &\longrightarrow X_{\phi_w}[0] \\
X[k_1] &\longrightarrow X_{\phi_w}[1] \\
X[k_2] &\longrightarrow X_{\phi_w}[2] \\
&\vdots \\
X[k_{\frac{P-1}{2}}] &\longrightarrow X_{\phi_w}[\tfrac{P-1}{2}] \\
X[k_{-\frac{P-1}{2}}] &\longrightarrow X_{\phi_w}[\tfrac{P+1}{2}] \\
&\vdots \\
X[k_{-2}] &\longrightarrow X_{\phi_w}[P-2] \\
X[k_{-1}] &\longrightarrow X_{\phi_w}[P-1]
\end{aligned} \tag{4.2}$$

Of course, conjugate symmetry should be exploited when recovering real signals.

The IDFT of the composite spectrum $X_{\phi_w}[j]$ yields the recovered waveform $x_{\phi_w}[m]$, the SPEC-PEAKS equivalent of the Rader algorithm composite period $x_{\tau_w}[m]$. If desired, $X_{\phi_w}[j]$ can be padded with an arbitrary number of zero samples inserted between the samples at $j = (P-1)/2$ and $j = (P+1)/2$, prior to inverse transforming.

We now consider the more common case where the signal fundamental frequency ϕ_w is not known. We begin by briefly reviewing the theoretical, iterative procedure

[3]We use the index j here to distinguish it from our other index k since the frequency scaling of the composite spectrum $X_{\phi_w}[j]$ and the aliased spectrum $X[k]$ will differ.

outlined in Chapter 2 (Figure 2.1).

The exact locations of the aliased harmonics were to be used as trial frequencies ϕ_g, or guesses of ϕ_w. Knowledge of the signal high-frequency cutoff ϕ_h $(= \Omega_h/\Omega_s)$ was required, though the allowable range of values for Ω_h was completely independent of the sampling rate Ω_s. Each ϕ_g was to be used in computing a tally of the number of its multiples[4] less than ϕ_h at which the aliased spectrum was non-zero. The true signal frequency ϕ_w would be given by the particular ϕ_g yielding the greatest tally.

For the moment, we assume that a set of trial frequencies ϕ_g including ϕ_w can be found. Nonetheless, we cannot collect an infinite number of samples, nor can we compute an infinite length DFT, so the true line spectrum of an aliased signal cannot be obtained. Therefore, we cannot compute meaningful tallies as in the ideal procedure from Chapter 2. The procedure must be modified in order to yield a practical algorithm.

Suppose that each ϕ_g is used to compute a corresponding *partial energy* $\mathcal{E}(\phi_g)$, which we define as the total spectral energy at all non-zero multiples of ϕ_g whose absolute values lie below some cutoff frequency ϕ_h:

$$\begin{aligned} \mathcal{E}(\phi_g) &= \sum_{\substack{i=L_1 \\ i\neq 0}}^{L_2} |X[k_i]|^2 \\ k_i &= \lfloor R\langle i\phi_g\rangle_1 \rceil \end{aligned} \tag{4.3}$$

The limits L_1 and L_2 are given by

$$L_2 = -L_1 = \left\lfloor \frac{\phi_h}{\phi_g} \right\rfloor \tag{4.4}$$

The partial energies can replace the tallies in the theoretical procedure from Chapter 2. They will serve the same purpose as the variations $\mathcal{V}(\tau_g)$ in the Rader algorithm, viz., to indicate the best estimate of the signal fundamental frequency. $\mathcal{E}(\phi_g)$ should be maximized at $\phi_g = \phi_w$. Assuming this is true in general, we now state a criterion for selecting the "best" trial frequency ϕ_{best} from some suitably chosen set:

Criterion 4.1 *The trial frequency which yields the greatest partial energy is the correct fundamental frequency, and the resulting waveform is the correct waveform.*

[4]Positive and negative multiples would be used for non-analytic signals, and their absolute values would be compared with ϕ_h.

We will refer to this statement as the *maximum partial energy criterion.*

Trial frequencies can be obtained by peak-picking the aliased spectrum $X[k]$ over the range between the known[5] limits

$$k_{min} = \lfloor R\langle \phi_{min} \rangle_1 \rceil$$

and

$$k_{max} = \lfloor R\langle \phi_{max} \rangle_1 \rceil$$

The spectrum must be sampled finely since the trial frequencies are now given by integers k_g which should correspond closely to the desired continuous frequency locations of the harmonics. We will see that, as was true in Chapter 2, using "baseband" trial frequencies k_g (i.e., values corresponding to residues modulo the sampling rate) will not affect the reconstruction process. Other issues related to the choice of trial frequencies, such as spectral sample spacing and number of trial frequencies, will be discussed in greater detail below.

We now have specified the basic framework of the SPEC-PEAKS algorithm: a procedure for choosing trial frequencies, a criterion for selecting the one which is the best approximation of the true fundamental frequency, and a procedure for reconstructing the signal using this estimate. However, there are three practical matters which still must be considered in implementing the maximum partial energy criterion.

First, due to spectral leakage, the collective peaks in the observed spectrum do not consist solely of aliased harmonics, as in the true spectrum. It follows that using all of these peaks as trial frequencies may produce erroneous results. For example, a leakage peak occurring at a submultiple of ϕ_w would surely yield a higher partial energy than the desired peak at ϕ_w itself would, since all energy in the latter case would have to be included in the former. If this frequency was chosen as the best estimate of ϕ_w, the reconstructed waveform period would contain several periods of the true signal. Due to the nature of the algorithm to be specified below, the waveform probably would lack more high-frequency information and certainly would contain more energy from spectral leakage than would a waveform reconstructed using ϕ_w. We therefore must

[5] As required by the pseudo-Nyquist criterion.

choose some subset of all the observed spectral peaks. It will be necessary to estimate the minimum number of trial frequencies needed to insure proper reconstruction.

Second, in the procedure in Figure 2.1, the energy summing process was terminated when the next computed multiple of the current ϕ_g was greater than some cutoff ϕ_h. The choice of ϕ_h was completely arbitrary in the sense that "excessively" large values would have no effect whatsoever on the tallies obtained. The true spectrum would be zero except at harmonic locations. However, this is not true for the observed spectrum, and thus all partial energies will include some spectral leakage components from the highest multiples of ϕ_g. Therefore, it is desirable to terminate the partial energy summation as soon as possible. We should either use a small ϕ_h, or choose another termination condition.

The last problem is due to the fact that the locations of the aliased harmonics can be determined only approximately from the sampled spectrum. Therefore, the partial energies computed using the resultant trial frequencies k_g will be slightly incorrect. In addition, even if the trial frequency closest to ϕ_w is chosen as k_{best}, and it does correspond to the index of the spectral sample nearest ϕ_w, the estimated harmonic amplitudes found using multiples of this value (just prior to reconstruction) will probably differ from their true values. Since this error increases in proportion to the original (analog) harmonic number, it may be desirable to omit high frequency harmonics in reconstructing the waveform. The IDFT of the composite spectrum would thus correspond to a low-pass filtered version of the original signal.

For simplicity, we require that P, the number of significant harmonics in the original signal, be known. We define significant harmonics as *all* analog harmonics in some frequency range $-\Omega_o < \Omega < \Omega_o$ (regardless of amplitude) which includes every harmonic whose amplitude is greater than some arbitrarily chosen minimum. Therefore, the number of significant harmonics at either positive (non-d.c.) or negative frequencies is given by

$$M = \frac{P-1}{2} \tag{4.5}$$

We now propose that the number of significant harmonics be used in all of the following cases, each intended to alleviate one of the three problems above:

1. Use M as the number of trial frequencies k_g.

2. Use M as the number of pairs of positive and negative multiples of each k_g to sum in computing the partial energy $\mathcal{E}(k_g)$.

3. Use P as the number of harmonics (including d.c.) in the composite spectrum for waveform reconstruction.

The reasons for these choices follow.

In general, there are equal numbers (M) of aliased significant harmonics in the lower and upper halves of the aliased spectrum, i.e., in the ranges $0 < k < R/2$ and $R/2 < k < R$. Since one of these two ranges must bound the baseband search range $[k_{min}, k_{max}]$, M is appropriate as the limit on the number of trial frequencies. The algorithm will fail if the spectral sample nearest the true fundamental frequency is not among the M largest spectral peaks in the search range (or if it is not a peak at all).

With regard to the second case enumerated above, we need to change the limits L_1 and L_2 in Equation 4.3 so that when the correct trial frequency is used, the corresponding partial energy contains no leakage components. The new limits are given by

$$L_2 = -L_1 = M = \frac{P-1}{2} \tag{4.6}$$

for non-analytic signals, and by

$$\begin{aligned} L_1 &= 1 \\ L_2 &= M = P - 1 \end{aligned} \tag{4.7}$$

for analytic signals. Beginning with Equation 4.2, we have assumed equal numbers of positive and negative harmonics, plus the one at d.c., for non-analytic signals. We will continue to concentrate our attention on non-analytic signals since in practice they are far more common, and the algorithm modifications needed for analytic signals are minor.

Using P as the number of harmonics in the composite spectrum is consistent with our choice of M as the number of pairs of positive and negative multiples of each trial frequency to sum in computing the corresponding partial energy.

The aliased input data $x[n]$ should be weighted with some windowing function $h[n]$ and then zero-padded before the DFT is computed. The choice of window type and size, as well as the total DFT length, has a significant impact on the quality of signal reconstruction. From our standpoint, the maximum amplitude of the window spectral sidelobes is most important. The signal cannot be recovered with the proposed algorithm if the fundamental is masked by leakage. A hamming window is a reasonable choice.

It is advantageous to use as many input samples as are available[6] since doing so decreases the chance of observed harmonics overlapping. On the ϕ frequency axis we have been using, the width of a hamming window main lobe is $4/N$, where N is the number of samples to be used, and therefore the length of the window, as well. When using such a window, the widths of observed aliased harmonics are approximately equal to this value.

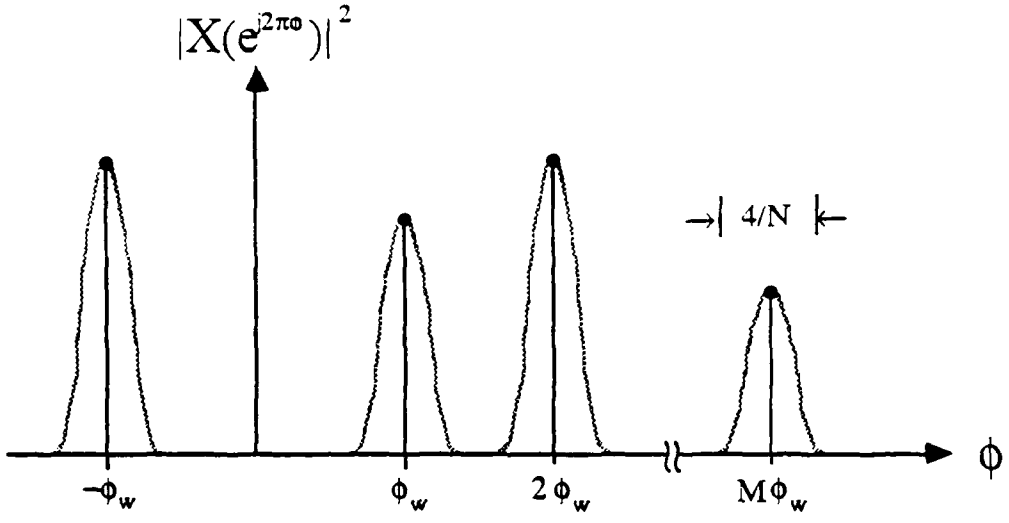
There is another important implication of the data/window size N. The DFT length R is also involved. Consider the effect of quantization along the frequency axis on partial energy computation. Figures 4.1(a) and 4.1(b) compare the ideal and non-ideal locations of the first few ϕ_g multiples used in computing $\mathcal{E}(\phi_g)|_{\phi_w}$. In the first case, $\phi_g = \phi_w$ exactly, while in the second, $\phi_g = \phi_w + \Delta\phi$. The frequency uncertainty $\Delta\phi$, due to quantization in ϕ, is bounded by the spectral sample spacing:

$$\Delta\phi \leq \frac{1}{2R} \tag{4.8}$$

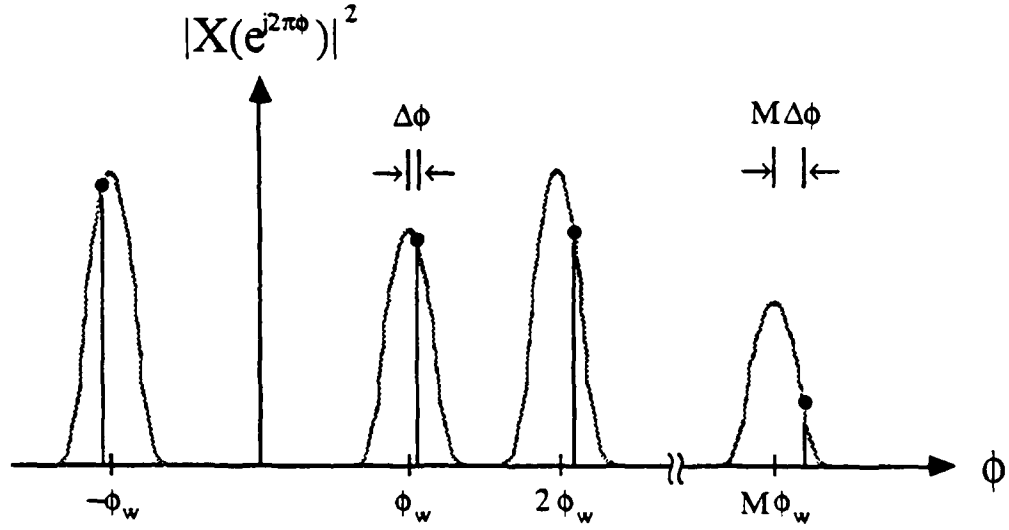
We now propose that the windowed sequence $x'[n] = x[n]h[n]$ be padded with enough zeros so that the value of $\mathcal{E}(\phi_g)|_{\phi_w}$ computed with the quantized ϕ_w will consist exclusively of energy from the smeared harmonics, as shown in Figure 4.1(b). This is equivalent to requiring that each multiple of the quantized ϕ_w correspond to a frequency somewhere on the appropriate smeared harmonic. Ideally, these multiples would correspond to the observed harmonic peaks, i.e., to the true harmonic locations.

If there are P significant harmonics, with equal quantities (M in Equation 4.5)

[6] With the possible limitation of maximum tolerable processing time.



(a)



(b)

Figure 4.1: First few trial frequency multiples used in computation of partial energy $\mathcal{E}(\phi_g)$. Images of harmonics due to aliasing not shown, for clarity. (a) Ideal case: $\phi_g = \phi_w$ (no quantization in ϕ). Hamming window width also indicated. (b) Non-ideal case: $\phi_g = \phi_w + \Delta\phi$.

at positive and negative frequencies, then the greatest error occurs in estimating the locations (and therefore, the amplitudes) of the M^{th} and $-M^{th}$ harmonics. From Equation 4.8, we see that the localization error of each of these two harmonics is less than $M/2R$. We desire that

$$M\Delta\phi \leq \frac{4/N}{2} \tag{4.9}$$

Therefore, the DFT length must satisfy

$$R \geq \left\lceil \frac{NM}{4} \right\rceil \tag{4.10}$$

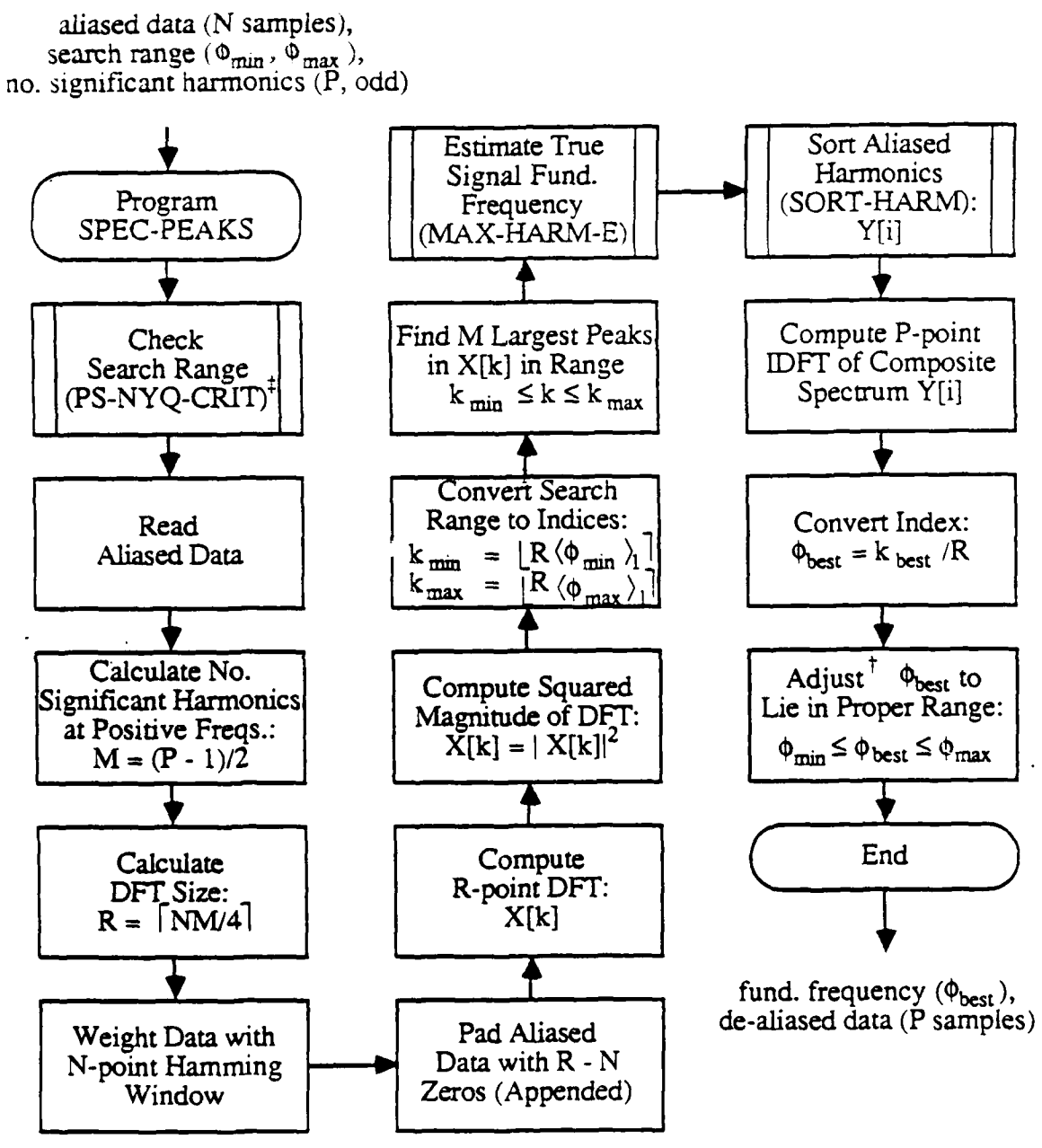
where $\lceil\ \rceil$ denotes the least higher integer or "ceiling" function. $R - N$ zeros must be appended to $x'[n]$.

Figures 4.2 through 4.4 contain flowcharts summarizing our implementation of the SPEC-PEAKS algorithm for non-analytic signals. They comprise a main program (SPEC-PEAKS) and three subroutines, all of which are called directly from the main program. Subroutine PS-NYQ-CRIT was shown previously in Figure 3.7.

Though the choice of the significant number of harmonics is clearly not well defined, the preceding algorithm has been used successfully in many instances, as will be shown in Section 4.3. We first will present a minor modification of the SPEC-PEAKS algorithm.

4.2 An Enhancement of the Algorithm

Towards the end of the previous section, we discussed the effects of quantization along the ϕ-axis on the computation of partial energies $\mathcal{E}(\phi_g)$. We are given N samples (or may elect to use only N samples when more are available) and assume the original signal contained P significant harmonics. We then choose the DFT size R so that when the location of the discrete spectrum peak nearest ϕ_w is used as a trial frequency ϕ_g, $\mathcal{E}(\phi_g)$ will contain energy from some portion of each smeared significant harmonic, and from no other regions of the aliased spectrum. Assuming the maximum partial energy criterion is correct, this peak location will be retained as the estimate of the true signal fundamental, ϕ_{best}.



Notes:
† Add appropriate multiple of ϕ_s (unity by definition).
‡ (See previous chapter.)

Figure 4.2: Program SPEC-PEAKS.

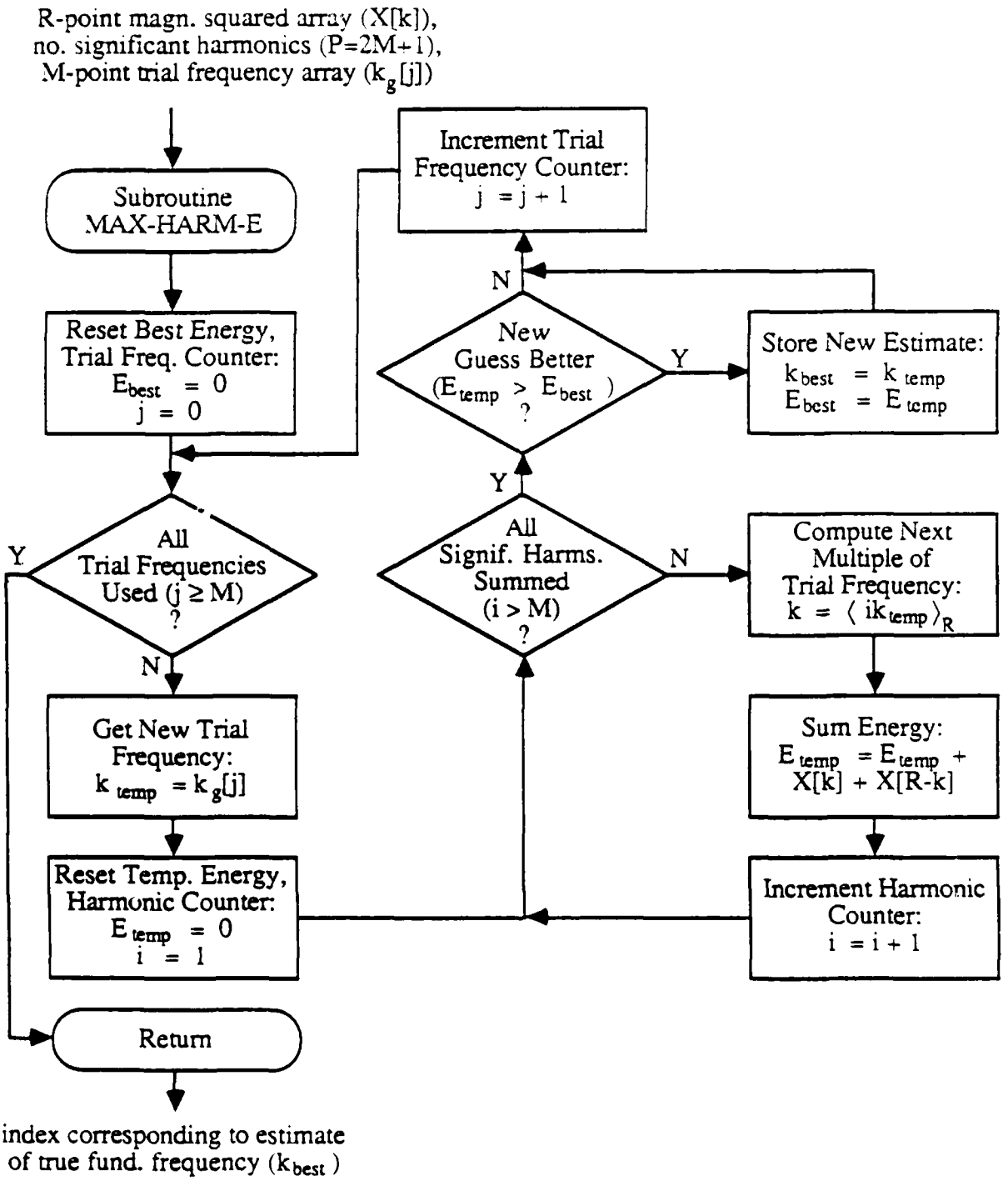


Figure 4.3: Subroutine MAX-HARM-E.

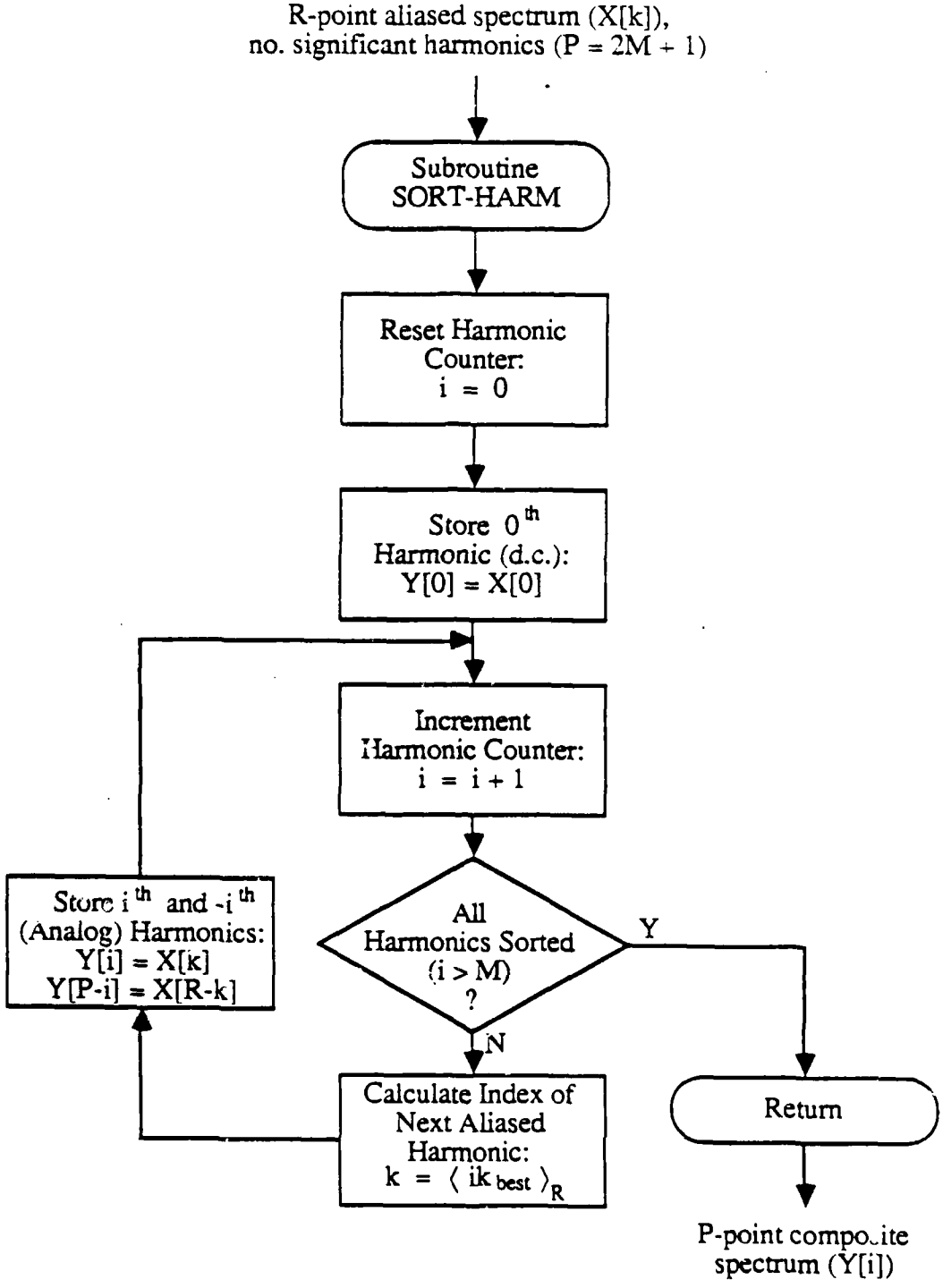


Figure 4.4: Subroutine SORT-HARM.

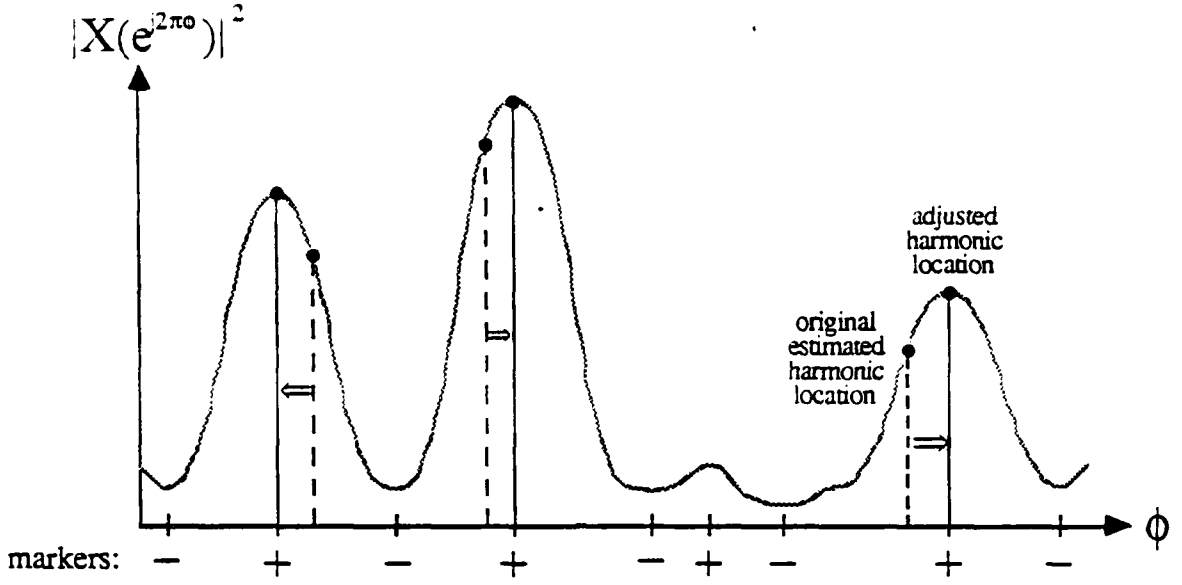


Figure 4.5: Adjustment of estimated harmonic locations for reconstruction.

In reconstructing the waveform with ϕ_{best}, the relative amplitudes of the sorted harmonics in the composite spectrum may deviate considerably from the original signal harmonic amplitudes. This is caused by error in aliased harmonic localization using multiples of ϕ_{best}, interpreted modulo ϕ_s.

If each estimated harmonic location lies on the correct smeared harmonic, this error can be reduced with the following procedure. Each estimated location is adjusted to the location of the peak of the lobe on which it lies,[7] just prior to forming the composite spectrum. The maxima and minima of the magnitude-squared spectrum are marked, as shown in Figure 4.5. The spectrum is then searched for the peak marker between the two minima markers that span the original estimated location. If a sufficiently large DFT size is used, and leakage ripples do not create extra peaks on the smeared harmonic main lobes, then the reconstructed waveform should more nearly resemble the original waveform.

The adjustment procedure can only be used in reconstruction. It cannot be used to increase the accuracy of the partial energies, all but one of which are computed using

[7] This does not imply the *nearest* peak.

an incorrect trial frequency.

A flowchart depicting the harmonic adjustment procedure is given in Figure 4.6. Also shown are the necessary modifications to program SPEC-PEAKS in Figure 4.2 and subroutine SORT-HARM in Figure 4.4.

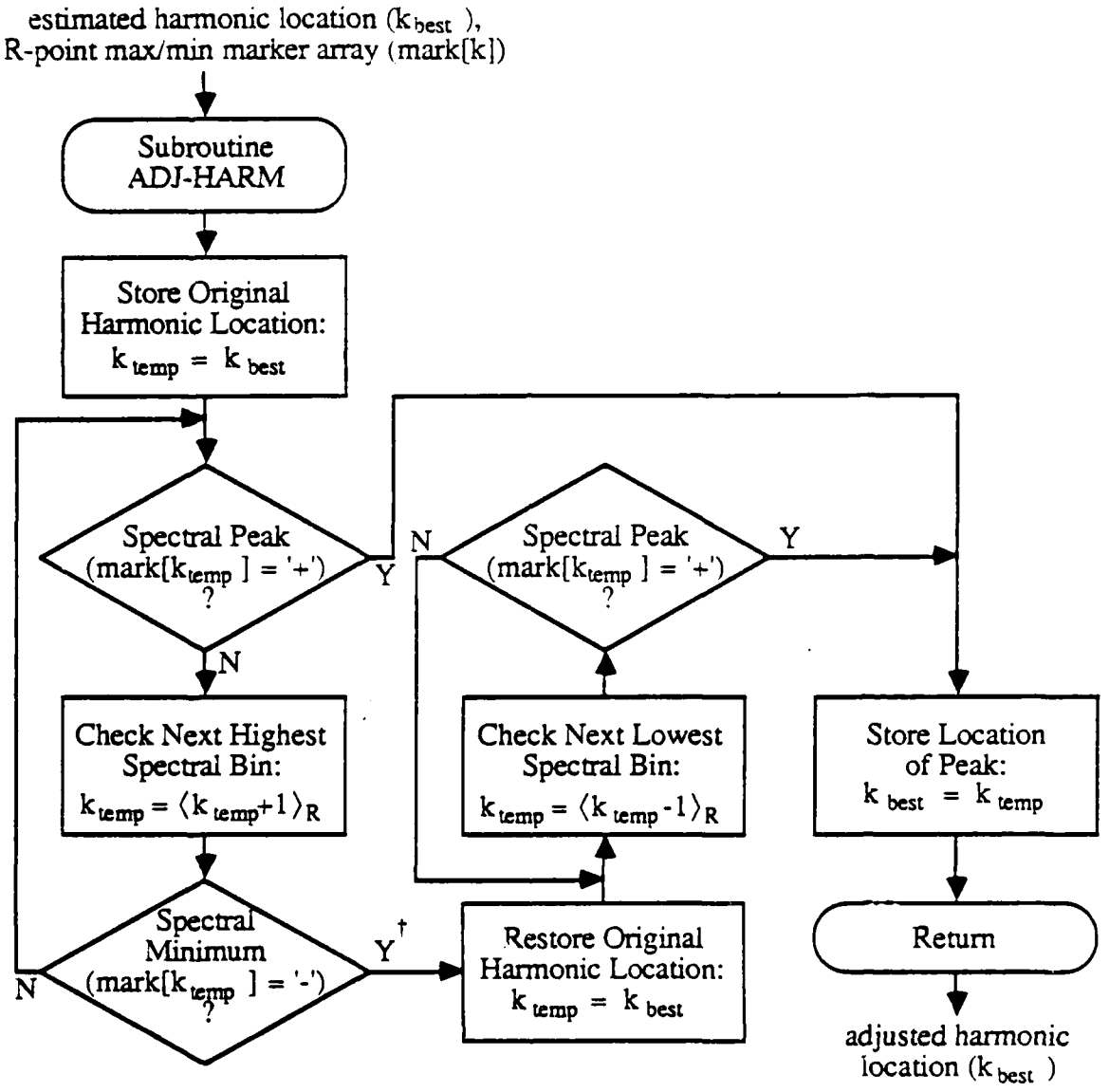
4.3 Examples

Figures 4.7 and 4.8 show examples of signals recovered with the SPEC-PEAKS algorithm, with and without the harmonic adjustment procedure described in Section 4.2. Since the same test signals were used for this chapter and the previous one, only the corresponding oversampled signals are repeated here, since they are useful for comparison with the recovered waveforms. The aliased signals can be found in Section 3.4 using the references provided in the new figures. Also noted in each caption is the normalized time scale factor by which the oversampled and recovered signals differ.

Figure 4.7(a) contains a test signal composed of ten equal-amplitude sinusoids superimposed on a small d.c. offset. The signal shown in Figure 4.7(b) was recovered with the unmodified SPEC-PEAKS algorithm. Figure 4.7(c) shows the signal recovered after harmonic adjustment. This waveform more nearly resembles the one in Figure 4.7(a). Less high frequency information has been lost, and the phase is correct. On the other hand, the error in the middle of the reconstructed period has been accentuated by harmonic adjustment. 1000 input samples were used for both cases, each requiring 7 seconds. (The harmonic adjustment time is negligible.)

Figure 4.8 contains waveforms originating from 60 Hz line interference sampled at 10796.123 Hz. The signal in Figure 4.8(b) was recovered with the unmodified algorithm. It is quite similar to the one in Figure 4.8(a), with the exception of the loss of a minute amount of high frequency information. As before, the third plot provides the recovered signal after harmonic adjustment. 1000 samples were used for each, requiring 8 and 9 seconds, respectively.

In both examples above, harmonic adjustment results in greater retention of higher harmonics since the estimated locations are moved to nearby spectral peaks (see Fig-



Notes:
† Reverse search direction.

(a)

Figure 4.6: Subroutine ADJ-HARM. (a) Subroutine. (b) Modification to program SPEC-PEAKS, Figure 4.2. (c) Modification to subroutine SORT-HARM, Figure 4.4.

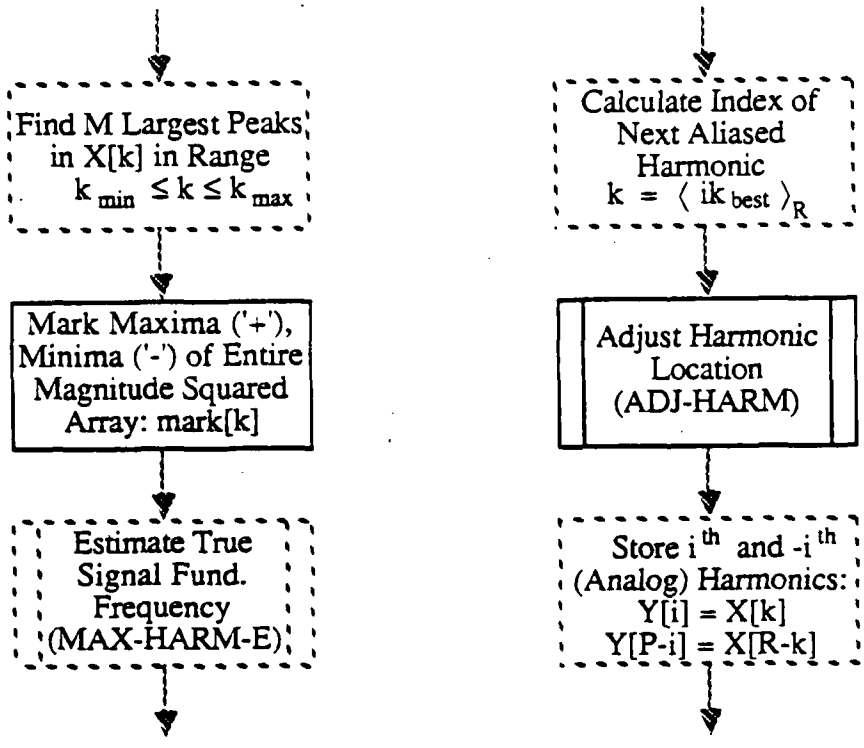


Figure 4.6: *continued*

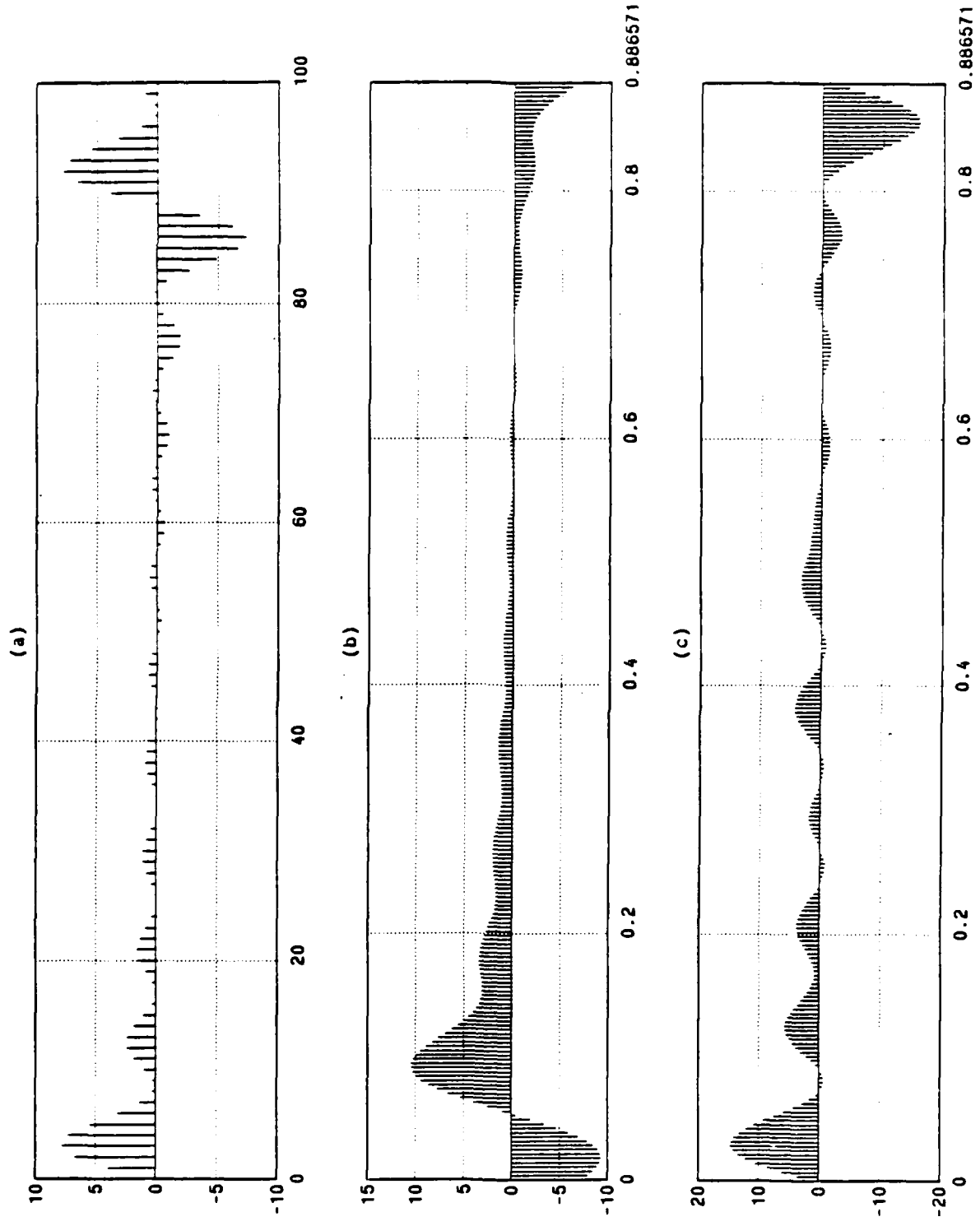


Figure 4.7: Aliased synthetic signal (ten sines with d.c. offset) recovered using SPEC-PEAKS. Compare with Figure 3.16. (a) Oversampled signal. (b) Recovered signal (time scale $1/100^{th}$ of that in (a)). (c) Recovered signal after harmonic adjustment (same time scale as (b)).

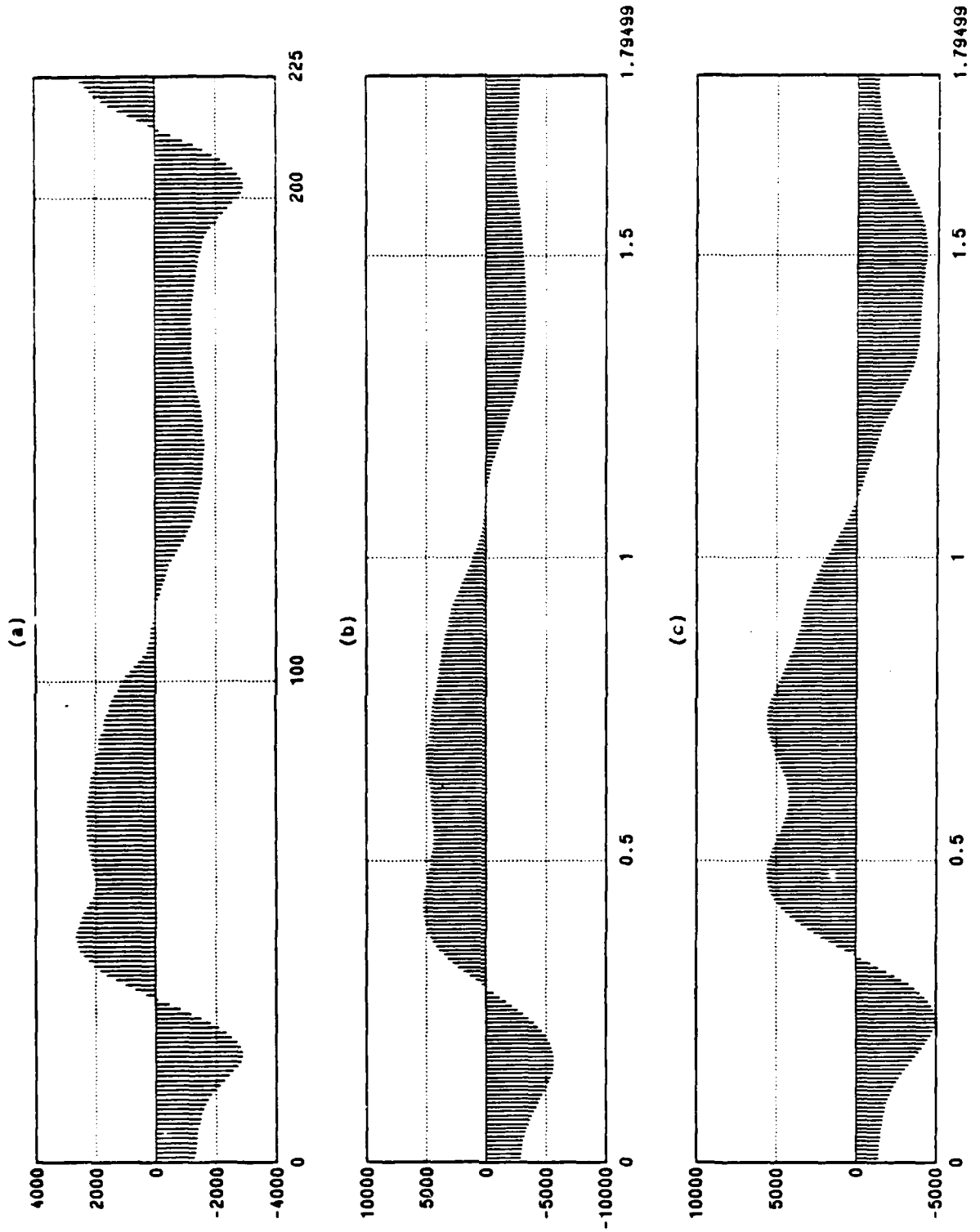


Figure 4.8: Aliased 60 Hz line interference signal recovered using SPEC-PEAKS. Compare with Figure 3.17. (a) Oversampled signal. (b) Recovered signal (time scale $1/100^{th}$ of that in (a)). (c) Recovered signal after harmonic adjustment (same time scale as (b)).

ure 4.5). However, the adjustment procedure clearly increases reconstruction error when the original harmonic locations lie on the wrong spectral peaks, as is the case for at least one harmonic in each of the two examples.

Chapter 5

Analysis and Conclusions

In this chapter, we will examine the accuracy and efficiency of the reconstruction algorithms described in the preceding chapters. There are four algorithms and variants to compare: the original Rader algorithm, with and without the FAST-SCAN modification entailing successively finer searches; and our SPEC-PEAKS algorithm, with and without the harmonic adjustment modification. Many of the trials will correspond to unfavorable conditions such as wavering signal amplitude and frequency, dynamic harmonic content, and additive noise. Representative output will be provided in the accompanying figures and tables.

Section 5.1 will discuss the quality of reconstruction achievable with the various algorithms. Particular attention will be paid to algorithm robustness, and cases likely to yield poor results. In Section 5.2, we will evaluate algorithm efficiency in terms of execution speed, and input and output data storage requirements. Since our harmonic adjustment modification affects only the accuracy (and not the speed) of our SPEC-PEAKS algorithm, it will only be treated in Section 5.1. Likewise, since our FAST-SCAN modification only increases the speed of the Rader algorithm, discussion of it will be limited to Section 5.2. Assuming the number of samples used[1] always exceeds the minimum number needed for sufficient reconstruction quality (see Section 3.3), the output from the unmodified and modified Rader algorithms will be identical.

[1] Per iteration in the FAST-SCAN case, and overall in the unmodified case.

5.1 Reconstruction Quality and Algorithm Robustness

In the course of algorithm development, many assumptions have been made, the validity of which can be ascertained only empirically. In this section, we will conduct tests which will indicate the classes of signals and types of conditions for which these assumptions fail. Plots of data from these tests will also permit subjective comparisons of the various algorithms.

Most plots will contain segments of the oversampled[2] and aliased sequences. Segments of the former typically will correspond to one or two periods of the original waveform; segments of the latter will represent portions of algorithm input corresponding to many periods. In contrast, the recovered waveforms will contain all the data used, and will always correspond to exactly one period, as evident from the algorithm descriptions presented earlier.

We know from Chapter 2 that any signal whose fundamental frequency is integrally related to the sampling rate generally cannot be recovered. Specifically, if the normalized signal period τ_w (or $1/\phi_w$, where ϕ_w is the normalized fundamental frequency) is a rational number u/v, then destructive aliasing (i.e., spectral overlap) may occur unless all signal harmonics occupy u or fewer adjacent harmonic locations. No algorithm, including the Rader algorithm and SPEC-PEAKS, can completely reconstruct such aliased signals. Nonetheless, it is interesting to examine distorted results.

Figure 5.1 contains reconstructions of an aliased, harmonically-rich test signal whose period τ_w is 10/7 (on the time scales in (b) through (d)). No matter how many aliased samples are collected, each corresponds to one of only 10 values, as apparent from the periodicity of the sequence in (b). This is also clear in Figure 5.1(c), in which the time samples have been sorted using the Rader algorithm. It might be possible to improve this reconstruction by discarding the redundant samples in this plot and interpolating the results, but this has not been investigated.

[2]The oversampled sequences will be shown only for comparison. In no case were they used as input to a reconstruction algorithm.

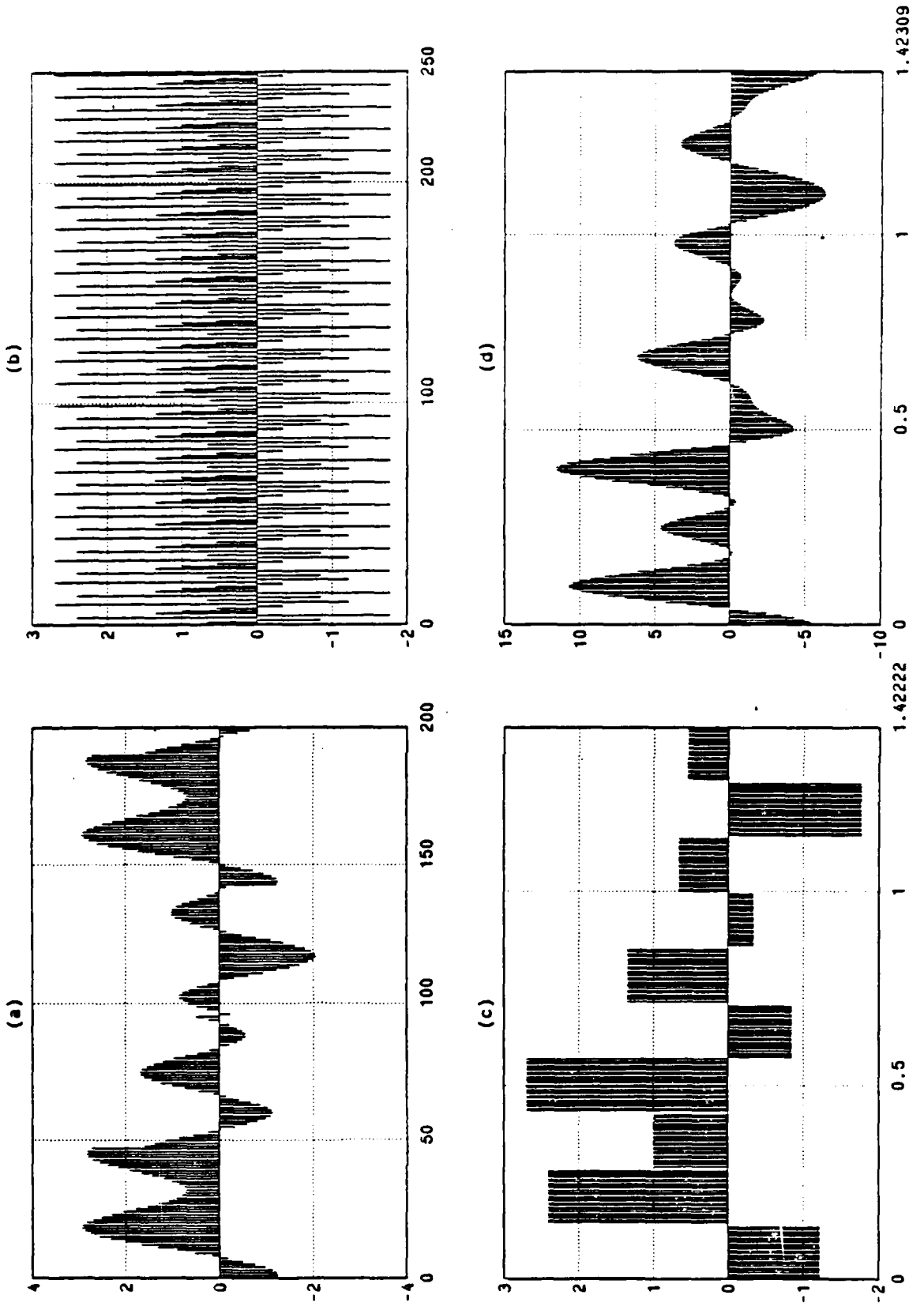


Figure 5.1: Poor reconstructions when pseudo-Nyquist criterion is not met. Test signal: three sines, one triangle wave. (a) Oversampled signal. (b) Aliased signal ((a) downsampled by 100). (c) Signal recovered using Rader algorithm. (d) Signal recovered using SPEC-PEAKS. (c) and (d) have same normalized time scale as (b).

Had the aliased signal been reconstructed with the Rader algorithm using the true signal period τ_w (which is itself a critical period) instead of the value found τ_{best}, considerable output sample overlap would have occurred. Since the Rader algorithm always returns an estimate τ_{best} which is not a critical period, this can never happen. In all cases similar to the present one, we would expect composite period samples with equal ordinates to be adjacent since τ_{best} and τ_w typically are nearly equal.

The corresponding results from the SPEC-PEAKS algorithm (Figure 5.1(d)) are also unacceptable. Note the presence of an unwanted peak between the two largest positive peaks, and the corresponding region in the Rader algorithm reconstruction. These effects were caused by the higher harmonics in the original signal being folded into the lowest ten harmonics.

Figure 5.2 presents an example of how the Rader algorithm excels in recovering aliased discontinuous waveforms. Provided that a waveform is smooth, with the exception of a few discontinuities, the minimum variation criterion would seem to hold. The results obtained using SPEC-PEAKS and assuming 40 significant harmonics are good, but a loss of high frequency information is evident. When the number of harmonics to be recovered was increased, the errors in localizing higher harmonics detracted from the quality of reconstruction.

The improvement obtained with the SPEC-PEAKS harmonic adjustment procedure is visible in comparing Figures 4.7(b) and (c) from Section 4.3. However, in many other cases this procedure proved to accentuate rather than reduce the error of SPEC-PEAKS in estimating harmonic amplitudes. In almost half the trials performed, the reconstructions were worse when harmonic adjustment was used.

The SPEC-PEAKS algorithm performs better than the Rader algorithm in dealiasing waveforms whose relative harmonic amplitudes change over the sampling interval, such as the one shown in Figure 5.3. This is due to the fact that the locations of the largest spectral peaks (i.e., the harmonic locations) typically remain unaffected when the complex harmonic amplitudes change. SPEC-PEAKS effectively integrates the information obtained over the sampling interval — the harmonic amplitudes in the reconstructed waveform represent average amplitudes.

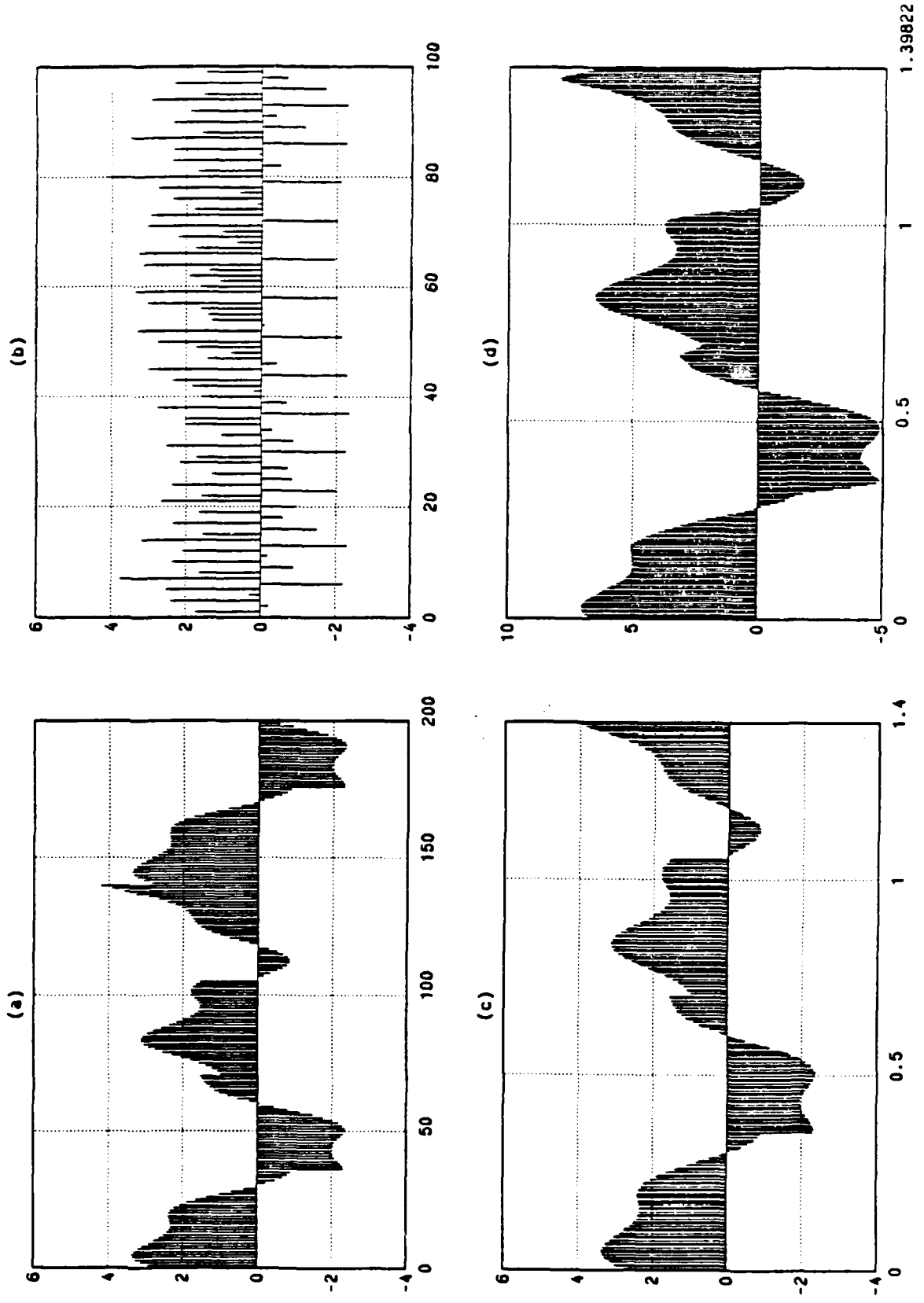


Figure 5.2: Reconstructions of discontinuous waveform. Test signal: three sines, two sawtooths, one square-wave. (a) Oversampled signal. (b) Aliased signal ((a) downsampled by 100). (c) Signal recovered using Rader algorithm. Note discontinuity at end of period ($\tau = 1.4$). (d) Signal recovered using SPEC-PEAKS. (c) and (d) have same normalized time scale as (b).

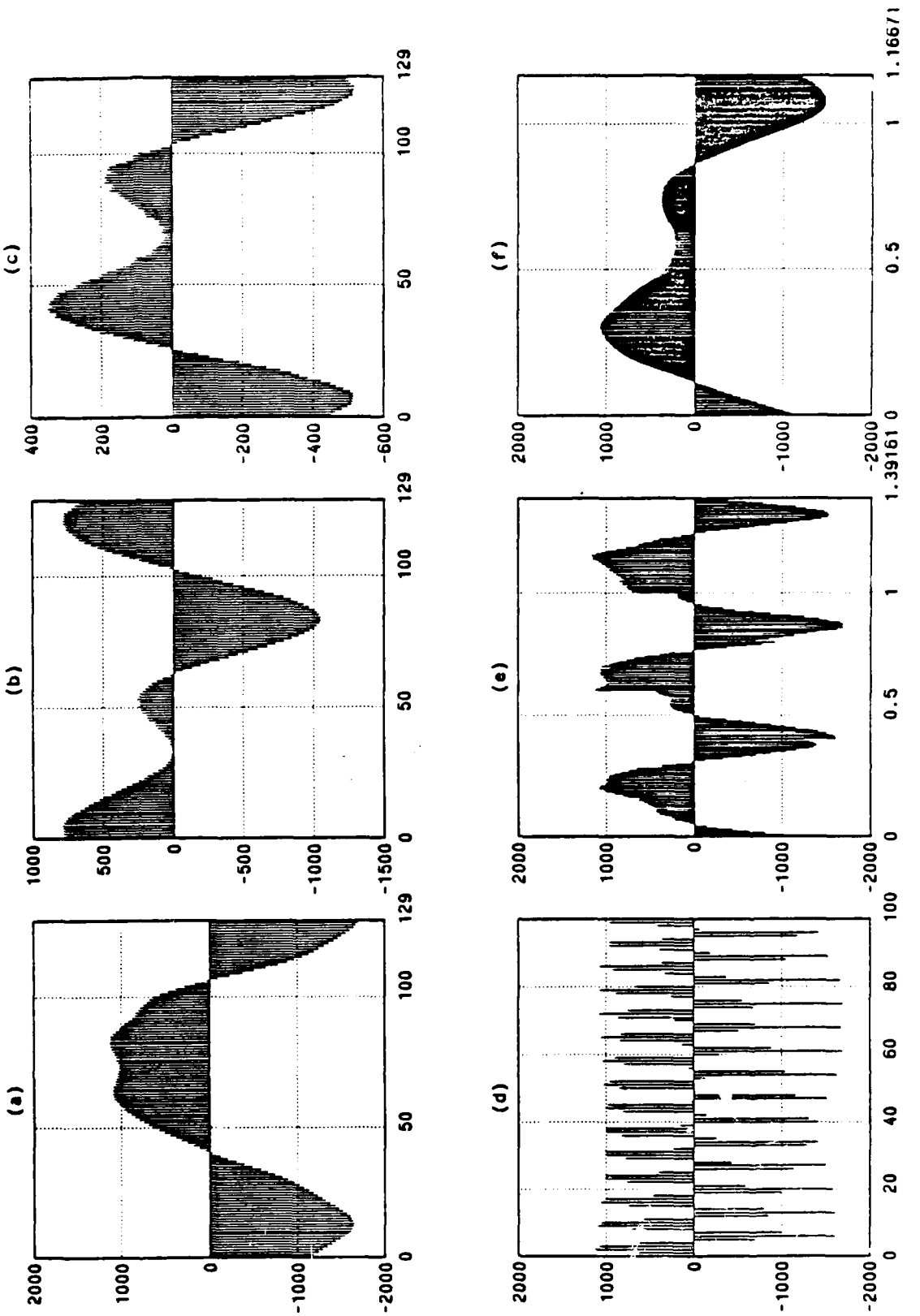
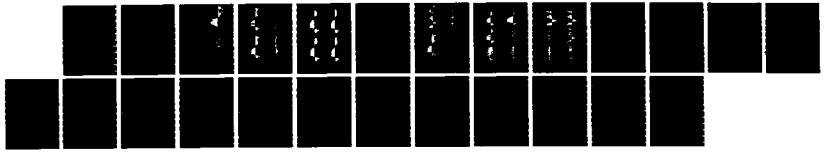
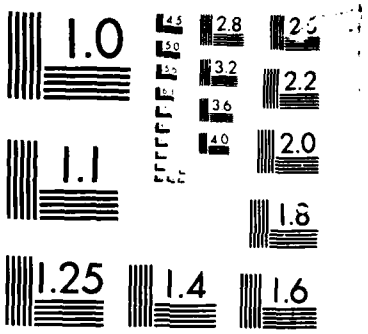


Figure 5.3: Reconstructions when relative harmonic amplitudes change. Test signal: guitar low E string (82.406889 Hz) sampled at 10 kHz. (a-c) Oversampled signal at beginning, middle, and end of sampling interval, respectively. (d) Aliased signal ((a) downsampled by 100). (e) Signal recovered using Rader algorithm. (f) Signal recovered using SPEC-PEAKS. (e) and (f) have same normalized time scale as (d).

AD-A168 897 RECONSTRUCTION OF UNDERSAMPLED PERIODIC SIGNALS(U) 2/2
MASSACHUSETTS INST OF TECH CAMBRIDGE RESEARCH LAB OF
ELECTRONICS A J SILVA JAN 86 TR-514 N00014-81-K-0742

UNCLASSIFIED F/G 9/3 NL




1.0
1.1
1.25
1.4
1.6
2.8
3.2
3.6
4.0
2.2
2.0
1.8

We see in Figure 5.3(e) that the Rader algorithm formed a composite period which appears to contain three similar periods. The width of the composite period (τ_{best} = 1.39161) is also incorrect. The minimum variation criterion does not hold here (nor was it intended to) since the ordinates of the original waveform at the same temporal positions within different periods are not equal, as they normally would be.

The performance of all algorithms is poor in recovering a waveform whose frequency or amplitude is not constant. The examples shown in Figure 5.4 correspond to a steady tone (the vowel *ä*) from a male speaker. Both the frequency and amplitude of the tone waver slightly over the sampling interval. The results from the Rader algorithm are similar to those in the previous example (Figure 5.3(e)). The images of several similar periods appear in the composite period (reversed in time, as well). Again, the minimum variation criterion does not hold for the same reason.

SPEC-PEAKS is only marginally better in that it is at least useful for determining the average fundamental frequency of the signal. The estimated frequency and average true frequency were very close. This is not an unreasonable result since the aliased harmonics are smeared by wavering frequency and amplitude, but the location of the main peak of each harmonic typically remains undisturbed. In both reconstructions, the original waveshape is distorted considerably.

Both the Rader and SPEC-PEAKS algorithms proved to be robust in terms of sensitivity to additive white gaussian noise. The four plots (a, c, e, and g) on the left side of Figure 5.5 correspond to a noiseless 60 Hz line interference test signal. Those on the right (b, d, f, and h) correspond to the test signal plus white gaussian noise. The signal-to-noise ratio (SNR) in Figures 5.5(b) and (d) is 18.4 dB.

Figure 5.5(f) shows the noisy aliased waveform in Figure 5.5(d) recovered with the Rader algorithm. The period determined here, τ_{best} = 1.79723, is identical to that in the noiseless case (Figure 5.5(e)). This is also true of the corresponding waveforms shown in Figure 5.5(g) and (h), recovered using SPEC-PEAKS. In both plots, τ_{best} = 1.79499. In addition, the SPEC-PEAKS algorithm removed nearly all the noise present in the aliased signal. However, since τ_w is known approximately, it seems reasonable that the Rader algorithm output can be processed with a comb filter to achieve results

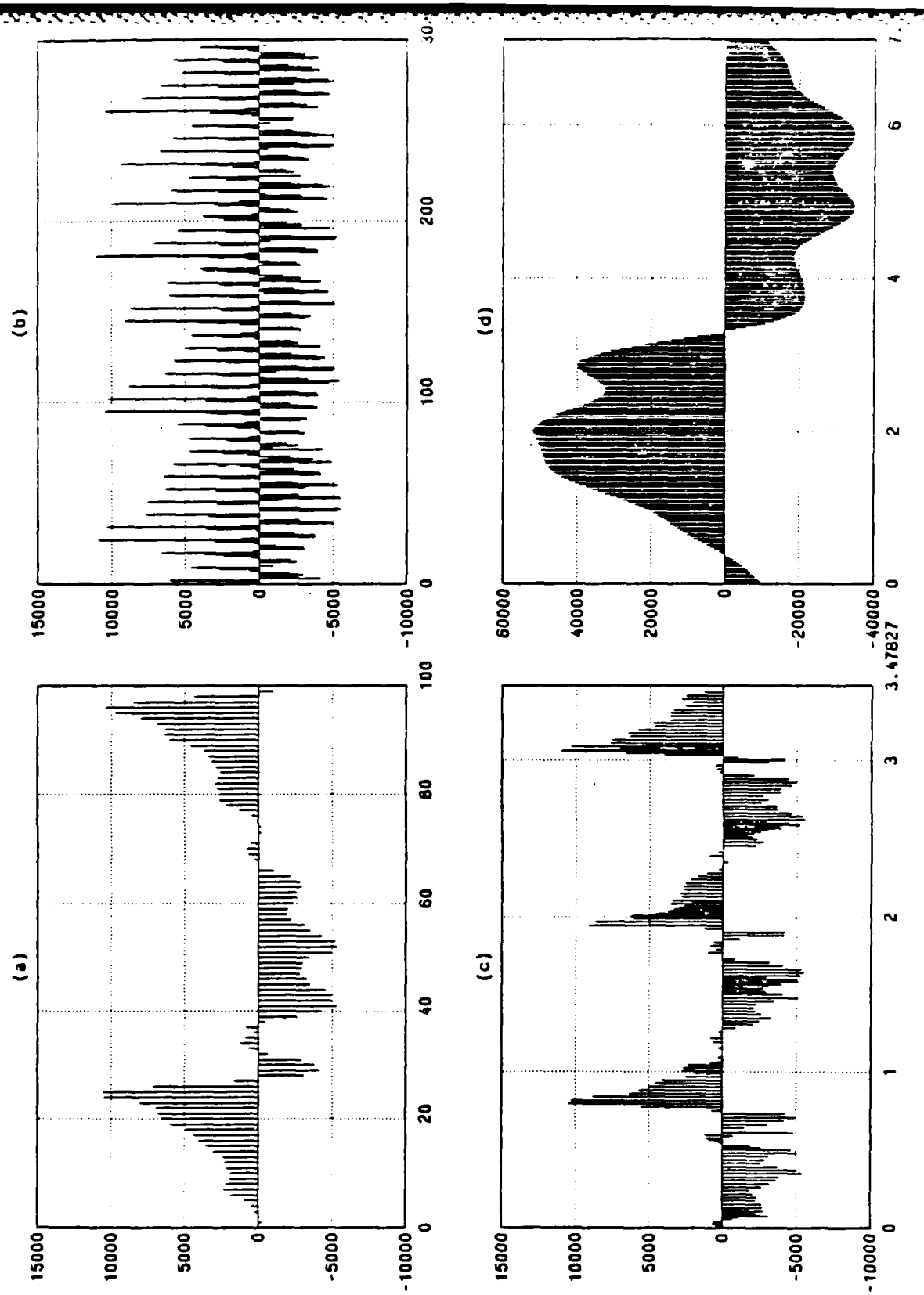


Figure 5.4: Poor reconstructions when fundamental frequency changes. Test signal: steady tone (vowel $\tilde{a}$) from male speaker, sampled at 10 kHz. (a) Oversampled signal. (b) Aliased signal ((a) downsampled by 10). (c) Signal recovered using Rader algorithm. (d) Signal recovered using SPEC-PEAKS. (c) and (d) have same normalized time scale as (b).

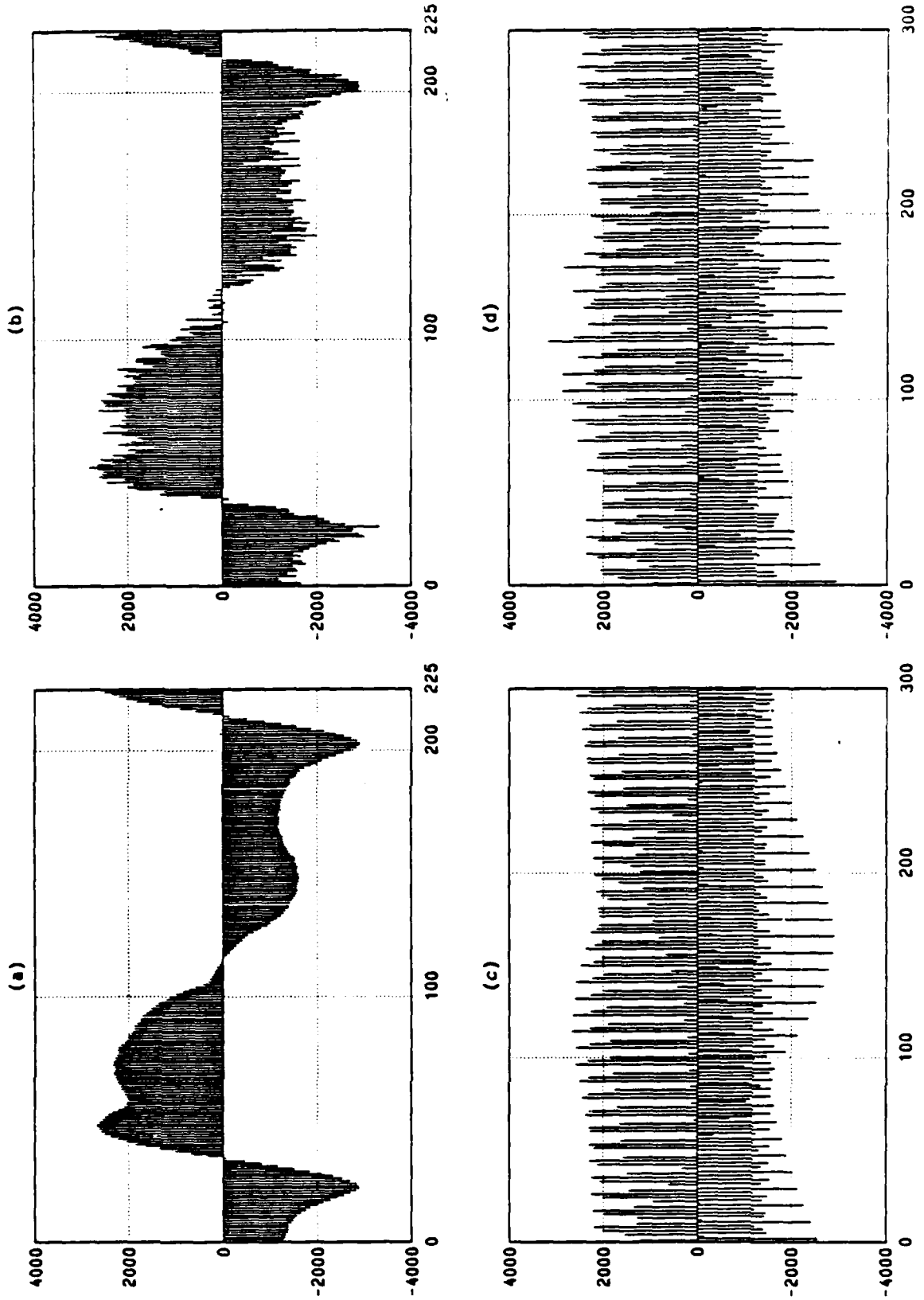


Figure 5.5: Reconstructions of noisy waveform. Test signal: 60 Hz line interference. (a,b) Oversampled signal with and without noise (SNR = 18.4 dB). (c,d) Aliased signal with and without noise (SNR = 18.4 dB). (e,f) Signals in (c) and (d) recovered using Rader algorithm. (g,h) Signals in (c) and (d) recovered using SPEC-PEAKS. Time scales in (c) through (h) $1/100^{th}$ of those in (a) and (b).

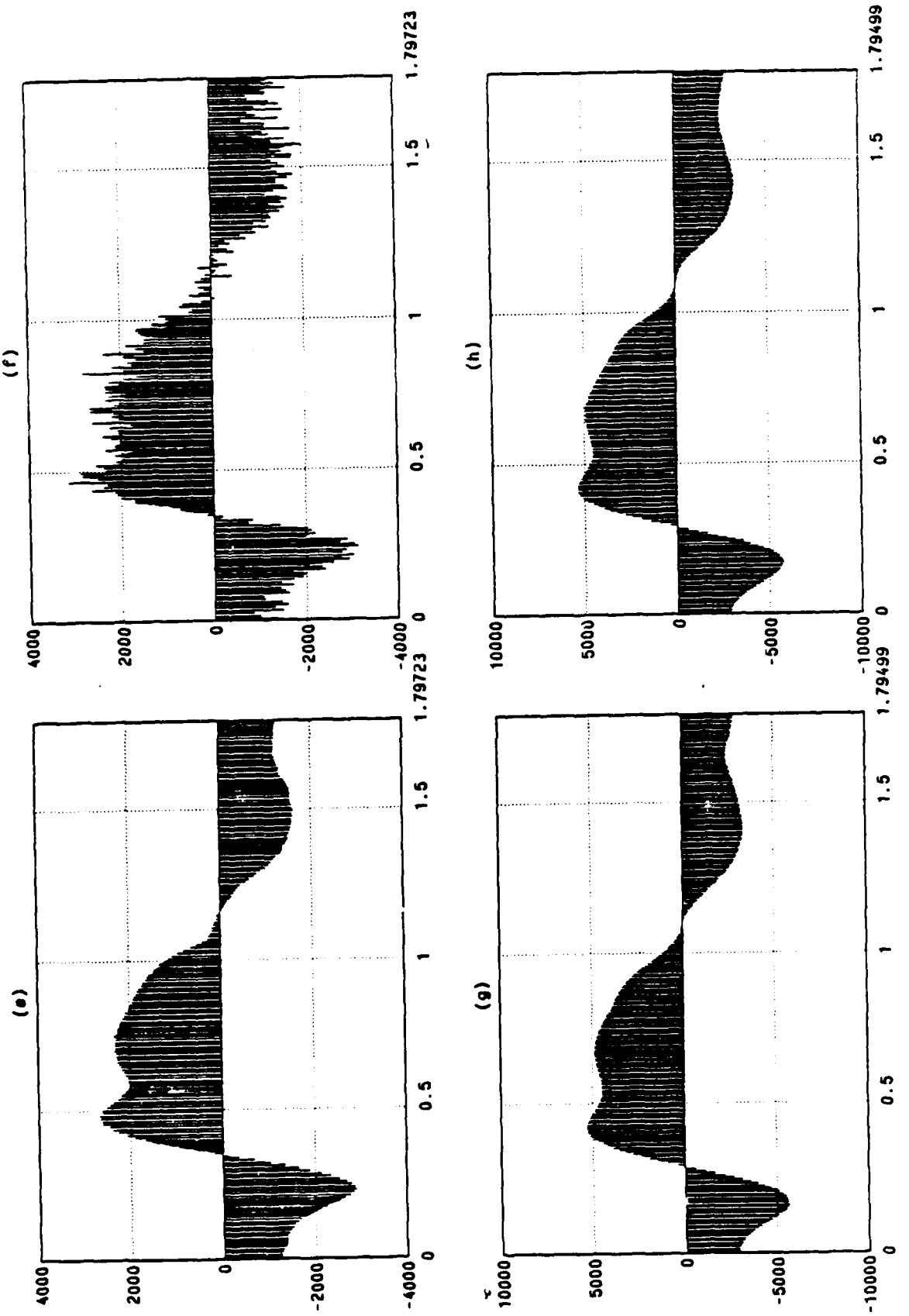


Figure 5.5: *continued*

comparable to those from SPEC-PEAKS. This has not been pursued further.

For the noise tests, the point of algorithm failure was defined as the SNR for which the estimated periods τ_{best} found in the noiseless and noisy cases (using a given algorithm) were no longer equal. The Rader algorithm did not fail until the SNR was decreased to 12.4 dB. SPEC-PEAKS was extremely robust, successfully estimating τ_w for SNR $= -9.53$ dB. However, all of these excellent results are surely due to the presence of a strong fundamental in the original signal.

Another additive corruption test was performed, consisting of the superposition of the two aliased test signals shown in Figures 5.6(c) and (d). For these tests, we defined a "signal-to-signal ratio"

$$\text{SSR} = 10 \log_{10} \left(\frac{\sigma^2_{sig1}}{\sigma^2_{sig2}} \right)$$

where σ^2_{sigQ} denotes the variance of signal Q.

The Rader and SPEC-PEAKS algorithms were used to recover the stronger of the two superimposed signals for various SSR, as summarized in Table 5.1. The definition of algorithm failure was analogous to that used previously in the noise tests. The weaker signal in each case was considered the corruptive one; thus, it replaced the noise in the definition above. By this criterion, SPEC-PEAKS was considered successful on every trial, as indicated by the boldface values in the table. The Rader algorithm failed to determine ϕ_w for SSRs in the range 10.1 to -19.4 dB.

Reconstructions from four of the trials enumerated in Table 5.1 are shown in Figures 5.6(e) through (p). These plots are arranged in groups of three (e.g., e, f, and g) comprising the aliased test signal and the waveforms recovered using the Rader and SPEC-PEAKS algorithms. The advantages of each algorithm are consistent with those seen in previous experiments: SPEC-PEAKS is much better at determining ϕ_w at low SSR (magnitudes), and in addition, removes most of the unwanted signal for higher SSR. On the other hand, the Rader algorithm retains more high frequency information in the dominant or recovered signal when successful.

We conclude this section with a few general reconstruction issues not discussed elsewhere. From Section 4.1 we know that the accuracy of the SPEC-PEAKS algorithm in determining ϕ_w depends directly on the DFT size. Therefore, the typical spectral

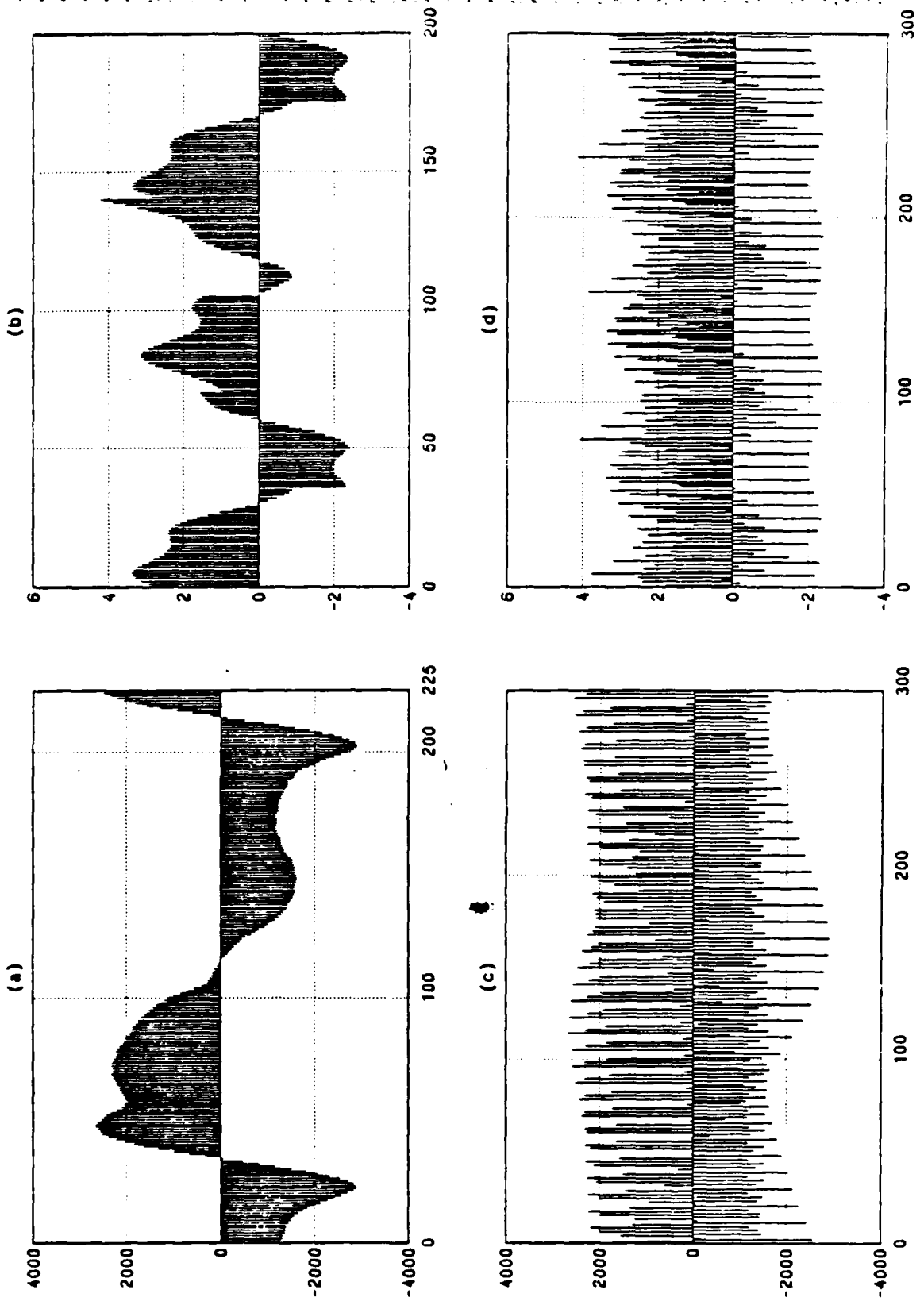


Figure 5.6: Reconstructions of two superimposed waveforms. (a) Test signal #1: 60 Hz line interference. (b) Test signal #2: three sines, two sawtooths, one square-wave. (c,d) Aliased test signals (#1 and #2 downsampled by 100). (e) Superimposed aliased test signals (SSR = 13.6 dB). (f) Signal in (e) recovered using Rader algorithm. (g) Signal in (e) recovered using SPEC-PEAKS. (h-j;k-m;n-p) Same as (e-g) for SSRs 0.55, −0.36, and −20.4 dB. Time scales in (c) through (p) $1/100^{th}$ of those in (a) and (b).

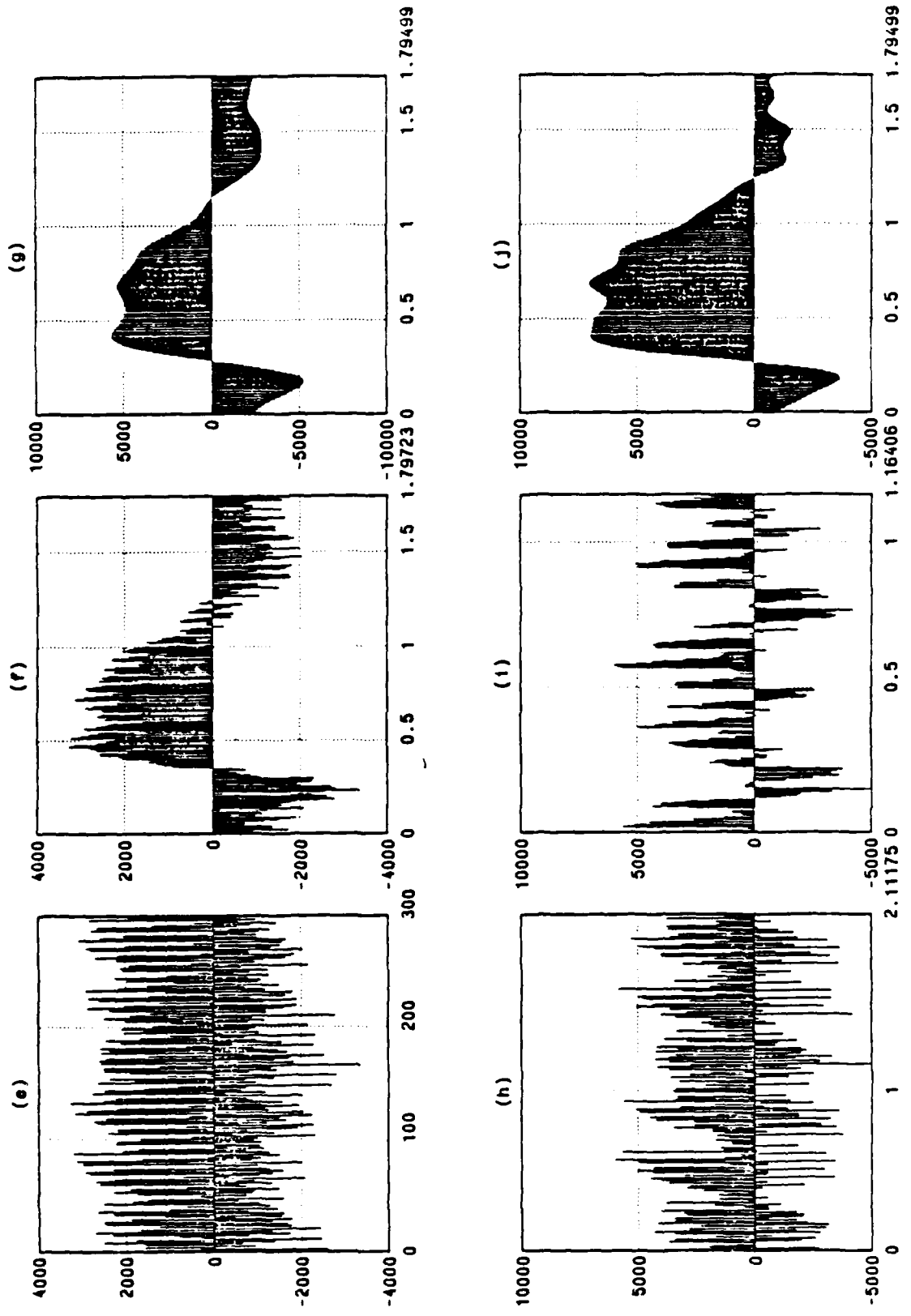


Figure 5.6: *continued*

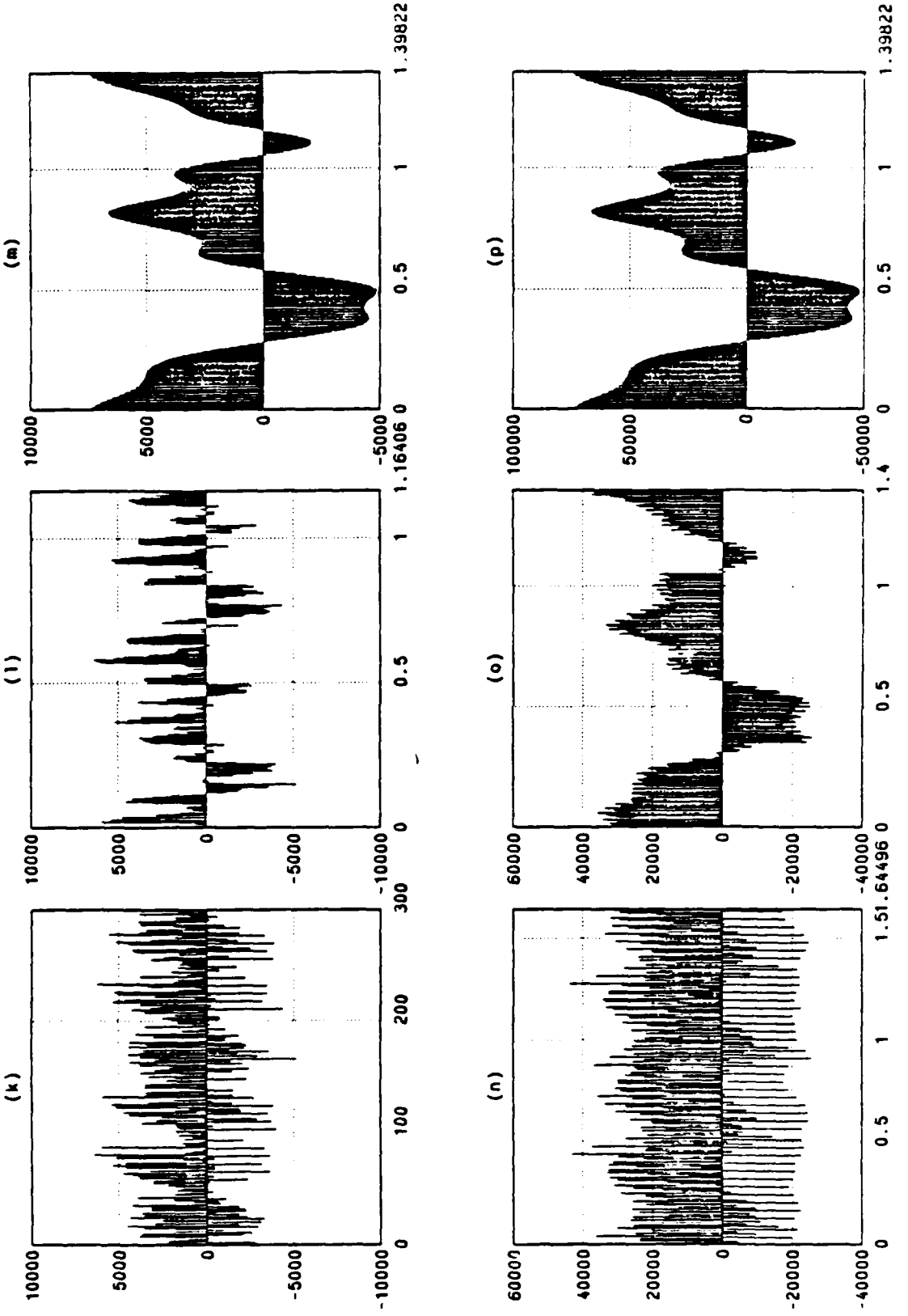


Figure 5.6: *continued*

SSR dB	ϕ_{best}	
	Rader algorithm	SPEC-PEAKS
19.6	**0.554987**	**0.554932**
13.6†	**0.554987**	**0.554932**
10.1	0.910979	**0.554932**
7.60	0.910979	**0.554932**
5.66	0.910979	**0.554932**
4.07	0.910979	**0.554932**
2.74	0.910979	**0.554932**
1.58	0.856187	**0.554932**
0.55†	0.856187	**0.554932**
−0.36†	0.856187	**0.712402**
−6.38	0.856187	**0.712402**
−9.90	0.856187	**0.712402**
−12.4	0.856187	**0.712402**
−14.3	0.856187	**0.712402**
−15.9	0.856230	**0.712402**
−17.3	0.856230	**0.712402**
−18.4	0.856230	**0.712402**
−19.4	0.856230	**0.712402**
−20.4†	**0.712435**	**0.712402**

Table 5.1: Estimation of ϕ_w from two superimposed waveforms. Test signal #1: 60 Hz line interference, $\phi_{w,sig1} = 0.555755$. Test signal #2: three sines, two sawtooths, one square-wave, $\phi_{w,sig2} = 0.712431$. Boldface numbers indicate results identical to those obtained from stronger aliased signal alone (using same algorithm). † signify trials plotted in Figure 5.6.

resolution is

$$R_s \approx \frac{2}{NM} \tag{5.1}$$

in units of normalized frequency, where N is the number of samples used, and M is the number of significant harmonics at positive frequencies.

The accuracy of the Rader algorithm is a function of the search density or Farey fraction order, which in turn depends on both N and the lower period search range limit τ_{min}. However, it can be estimated only roughly since the Farey fractions are distributed non-uniformly. Rader has stated that the average period accuracy can be obtained from the reciprocal of the density of the Farey fractions. For a given order L, there are approximately

$$\frac{3L^2}{\pi^2} + O\{L \log L\}$$

Farey fractions on any interval[3] $(r, r+1)$, where r is a real number. Since $L = (N-1)/\tau_{min}$, the average temporal resolution is

$$R_t \approx \frac{\pi^2 \tau_{min}^2}{3N^2} \tag{5.2}$$

in units of normalized time. Rader has also mentioned that in the worst case, R_t becomes τ_{min}/N, and in addition, both this and the estimate in Equation 5.2 are optimistic in that the density of the critical periods (which truly determines the search density) is less than the density of all the Farey fractions.

It would be desirable to avoid the $u - N$ missing samples in Rader algorithm reconstructions, where u is the numerator of the period τ_{best} returned. One might consider not using all the available data so that $u - N$ leftover samples could be inserted into the reconstructed period. However, this should not be done since these $u - N$ samples would have very erroneous ordinates for the missing samples they were intended to fill. Had we used u $(> N)$ samples in the first place, the search density would have been finer $(L = \lfloor (u-1)/\tau_{min} \rfloor)$, and there still would have been missing samples.

[3]The Farey fractions in any interval $(r+i, r+i+1)$, where r is a real number and i is an integer, can be obtained by adding i to each Farey fraction in $(r, r+1)$. Therefore, the average densities in any such pair of intervals are equal.

N	$t_{reconst}$ sec
25	3
50	7
100	35
200	120
300	300
500	1400
1000	11000

Table 5.2: Rader algorithm recovery time vs. number of input samples, N. Typical values shown. Search range: 0.5–1.0.

5.2 Algorithm Efficiency

As could be expected, the speed of all algorithms is dominated by N, the number of samples used. The speed of the Rader algorithm, like its accuracy, depends on the density of the critical periods in the search range. One value of the variation function is computed between each pair of successive critical periods, each requiring computation of order $\mathcal{O}\{N\}$. Therefore, from the previous section (Equation 5.2) we know that for a given search range, the total computation is typically $\mathcal{O}\{N^3\}$. In all cases, it is greater than $\mathcal{O}\{N^2\}$.

Typical reconstruction times versus number of input samples are given in Table 5.2. The values in this table, as well as the tables to follow, correspond to real time (i.e., elapsed, not cpu time). All timed experiments in this thesis were performed during periods in which system load averages were low (typically overnight, when there were no interactive users on the system).

Equation 5.2 indicates that the Rader algorithm speed also depends on the upper limit of the fundamental frequency search range, $\phi_{max} = 1/\tau_{min}$. We will now show that any search range ϕ'_{min}–ϕ'_{max} whose respective limits are congruent modulo one ("ϕ,") to the true limits ϕ_{min} and ϕ_{max} can be used. This allows us to pick a range expected to reduce execution time. We would then reconstruct the waveform *exactly* as before. Finally, we would relabel the time axis of the reconstructed period and temporally

reverse the data (if necessary) to yield the same results that would have been obtained with the original search range.

Recall from Chapter 2 that identical sequences are formed in sampling two analog signals whose normalized fundamental frequencies ϕ_{w1} and ϕ_{w2} differ by some integer. Alternately, if the sum of ϕ_{w1} and ϕ_{w2} is an integer, one sequence is the time reversal of the other. In fact, it is for these reasons that the correct waveshape can *always*[4] be recovered, even if the proper search range is unknown. In the latter case, any search range not spanning a multiple of $1/2$ could be used with either the Rader algorithm or SPEC-PEAKS.

Figure 5.7 contains two plots of the variation function produced while recovering the waveform shown in Figure 5.8(a) using the Rader algorithm. A search range ϕ_{min}–ϕ_{max} spanning several multiples of $1/2$ (thereby violating the pseudo-Nyquist criterion) was used. Plotted versus the reciprocals of the trial periods τ_g (Figure 5.7(b)), $\mathcal{V}(\tau_g)$ is periodic. Note the similarity between the portions of the variation function in the regions $\phi_g = 0.5$ to 1.0 and $\phi_g = 1.5$ to 2.0, as well as the symmetry about integral ϕ_g (in particular, $\phi_g = 1.0$). The presence of equivalent minima in these three regions clearly indicates the fundamental frequency ambiguity problem addressed in Chapter 2.

We now return to our examination of the effect of different search ranges on Rader algorithm execution time. The pseudo-Nyquist criterion dictates that each search range must lie between two successive multiples of $1/2$. Figures 5.8(b) through (h) contain reconstructions of the test signal in Figure 5.8(a) after it was downsampled by 100. The corresponding search ranges are listed in Table 5.3, along with the reconstruction times. The seven reconstructions are identical (with the exception of the time reversal in every other plot). However, the search times for search ranges $p/2$–$(p+1)/2$ (for integral p) decrease monotonically, as do those for ranges $(p+1)/2$–$p/2$. The decrease of the aggregate is almost monotonic, as well.

The latter result can be explained as follows. The same number of Farey fractions of a given order lie in each range $(i, i-1)$ along the τ_g-axis for all integers i. Therefore,

[4] Assuming no destructive aliasing has occurred.

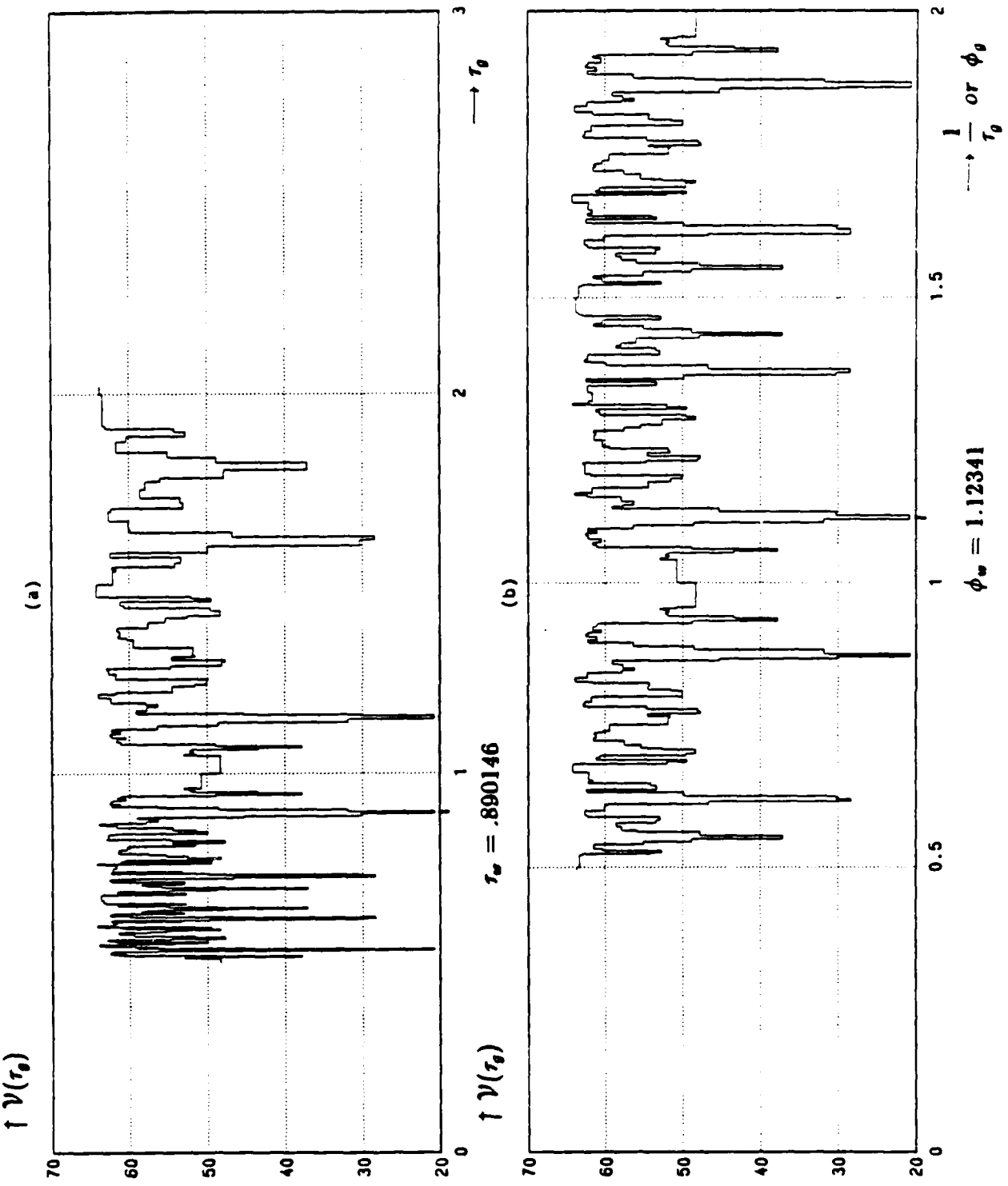


Figure 5.7: Periodicity of variation function when plotted vs. trial frequency. Test signal: ten sines with d.c. offset. (See Figure 5.8.) (a) Variation vs. trial period. (b) Variation vs. trial frequency.

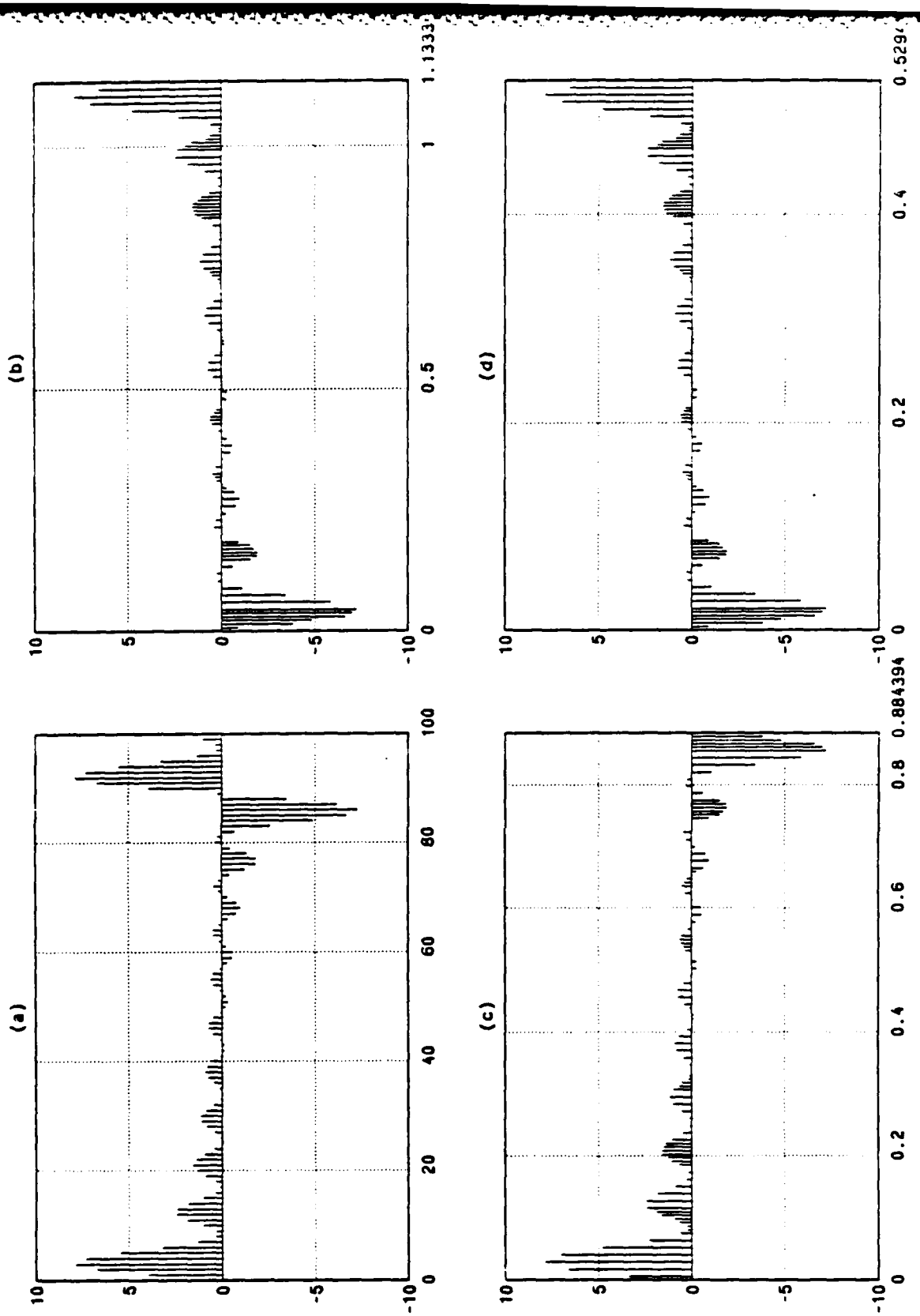


Figure 5.8: Rader algorithm reconstructions for same aliased signal using different search ranges. (a) Oversampled test signal: ten sines with d.c. offset. Time scale 100× that in other plots. (b-h) Signals recovered from (a) downsampled by 100 using Rader algorithm. Search ranges and numerical results listed in Table 5.3. Correct search range was used in (c).

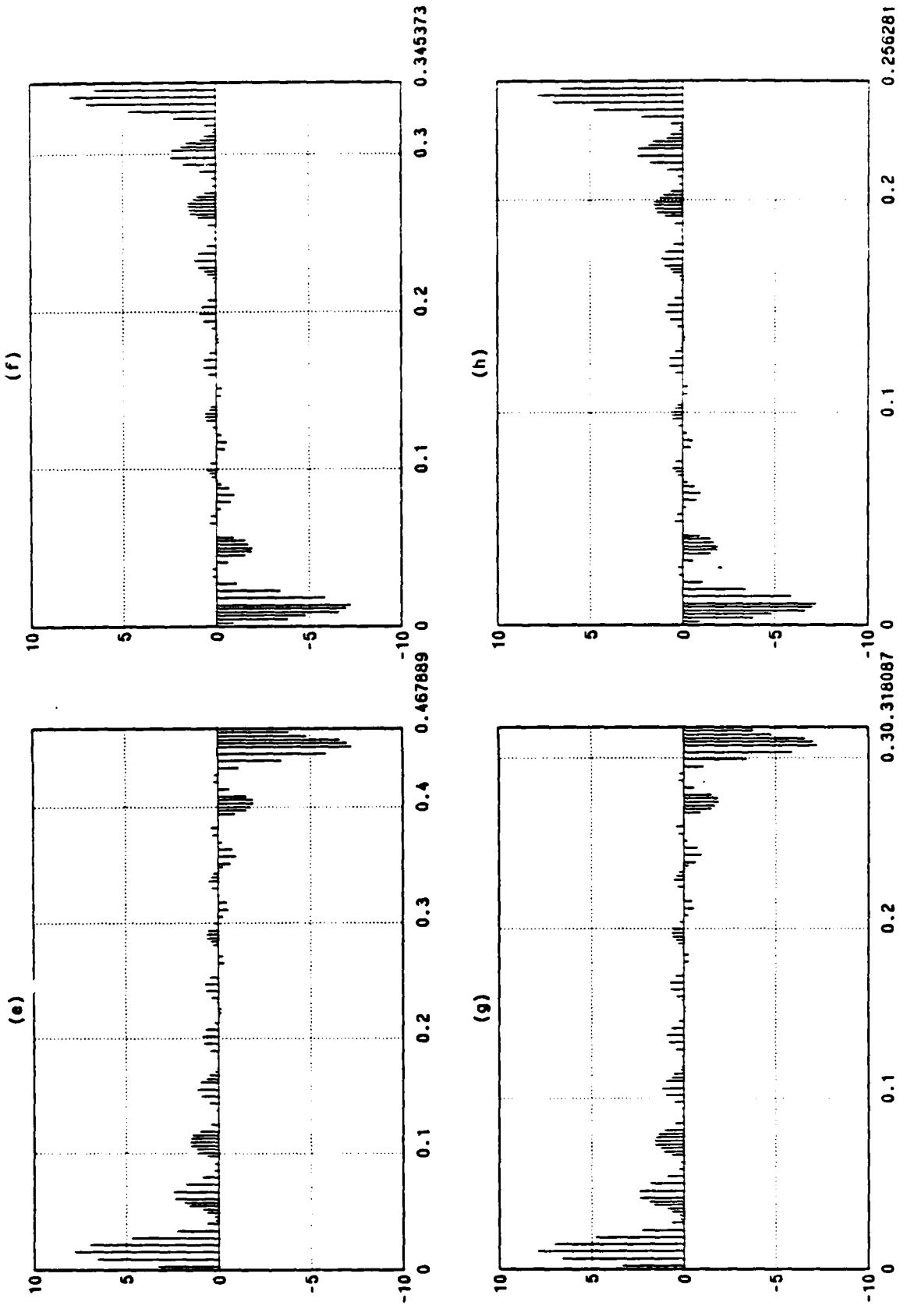


Figure 5.8: *continued*

ϕ_{min}–ϕ_{max}	ϕ_{found}	$t_{reconst}$ sec	figure
0.5–1.0	0.87662	47	b
1.0–1.5	1.12338	33	c
1.5–2.0	1.87662	29	d
2.0–2.5	2.12338	16	e
2.5–3.0	2.87662	21	f
3.0–3.5	3.12338	14	g
3.5–4.0	3.87662	16	h

Table 5.3: Rader algorithm search time vs. search range. Same number of samples from same aliased signal used in each case. References to Figure 5.8 also given. Correct search range: 1.0–1.5.

they become increasing sparse along the positive ϕ_g-axis since $\phi_g = 1/\tau_g$. On the other hand, a higher search range ϕ_{min}–ϕ_{max} (equivalently, $1/\tau_{max}$–$1/\tau_{min}$) causes the Rader algorithm to use a higher Farey fraction order, as indicated by Equation 3.9. The decrease in the search times given in Table 5.3 is due to the fact that these two effects do not cancel. The former seems to be slightly stronger. Therefore, the search density and execution time tend to decrease for higher search ranges.

While the reductions in Rader algorithm reconstruction time can be large[5] when using our FAST-SCAN modification, they cannot be quantified even roughly (empirically or otherwise) since the initial Farey fraction order and increment are chosen heuristically. The minimum initial order for which the minimum variation criterion holds cannot be determined.

The time required by SPEC-PEAKS is typically of order $O\{NM \log NM\}$ where M is the number of significant harmonics at positive frequencies, since most of the processing time is spent computing the finely sampled DFT. In practical situations, the size of the IDFT producing the output is much smaller. Recovering waveforms using 1000 samples and assuming 10 significant harmonics at positive frequencies typically required 5-11 seconds. An example of the dependence of SPEC-PEAKS on the number of harmonics is summarized in Table 5.4, in which 1000 samples were used on each

[5] An order of magnitude or more was not uncommon, but the comparision has little merit.

M	$t_{reconst}$ sec
10	7
20	12
30	14
40	25
50	26
60	27
70	53

Table 5.4: SPEC-PEAKS recovery time vs. number of significant harmonics at positive frequencies (M). Typical values shown. 1000 samples used.

trial. No other tables are given since the speed performance of SPEC-PEAKS versus N is predictable from the discussion above.

In contrast to the Rader algorithm, neither the accuracy nor the speed of SPEC-PEAKS are effected by the search range ϕ_{min}-ϕ_{max}. One of the first steps in the SPEC-PEAKS algorithm consists of converting the search range to the corresponding (baseband) DFT indices.

It was found that SPEC-PEAKS typically required an order of magnitude more data than the Rader algorithm to yield comparable reconstructions. SPEC-PEAKS was still much faster in these cases, but the comparison must be viewed in light of the fact that SPEC-PEAKS always loses some high frequency information. Another advantage of the SPEC-PEAKS algorithm is that it can place the output data in minimum storage form (i.e., sampled just above the Nyquist rate) with no additional computation. In such cases, the composite spectrum would not be zero-padded.

Chapter 6

Suggestions for Future Research

We conclude this thesis with a list of suggestions for future research.

The pseudo-Nyquist criterion we have developed may be too restrictive in certain cases (as was true of the original Nyquist criterion, as well). We have shown that the irrationality requirement is unnecessary if Ω_s/Ω_w can be expressed as a rational number u/v where $(u,v) = 1$, and fewer than u consecutive signal harmonics are present. This issue is coupled with the requirement for an arbitrary high frequency cutoff Ω_h since the latter insures a finite (though unknown) number of harmonics. Alternative definitions of the pseudo-Nyquist criterion are certainly plausible.

Most of the Rader algorithm processing time is used to compute the collective variation functions. Significant savings should be possible by exploiting the fact that only a few composite period samples interchange in moving from one trial period to the next. (Visualize the effect of increasing the diameter of the cylinder in Figure 3.1.) If the original indices n of these samples $x[n]$ can be determined, the variation from the previous iteration can be reused. A few terms would then be added to correct for the interchanging samples. Only one of the samples must be located since for a given trial period $\tau_g = u/v$, the multiplicative inverse of v for the modulus u can be used to locate the others (see Section 3.2). The fact that aliased samples $x[n]$ at higher indices n would move clockwise[1] in the composite period faster as τ_g is increased might also be useful here. Finally, if it appears that the overhead incurred is high (e.g., if many

[1] Viewing the cylinder in Figure 3.1 from the top.

samples interchange), the variation could be computed in the original manner.

Interpolation of missing samples in Rader algorithm reconstructions should be explored. Methods which are insensitive to sample spacing (e.g., Lagrange interpolation) would seem to be most suitable. Conventional discrete time filtering techniques may also be appropriate since the reconstruction samples (including those which are missing) lie on a uniformly spaced axis. Rader has suggested moving median filtering, followed by linear filtering. The median would be redefined as the mean of the medians of subgroups of points. Redefinition is necessary since the conventional median of three consecutive points is undefined if one or more of the three are missing.

Another interpolation method which may be applicable has been proposed by Naidu and Paramasivaiah [11]. It comprises an extension of the Gerchberg-Papoulis algorithm [12], originally for extrapolation of band-limited signals, to interpolation of missing samples. Knowledge of the bandwidth of the original signal and an average sampling rate[2] above the Nyquist rate are required, but neither condition should pose a problem here. Marks [13] has also succeeded in extending the Gerchberg-Papoulis extrapolation algorithm to missing sample interpolation.

The harmonic adjustment modification of the SPEC-PEAKS algorithm does improve reconstruction in many cases. However, it would be desirable to determine when this is not true so that harmonic localization errors would not be compounded by unwanted adjustments.

Rader [1] has suggested that his algorithm may be useful in simultaneously reconstructing several signals $x_{a,i}(t)$ with equal periods, each sampled at the same rate. The aliased sequences $x_i[n]$ could be treated as a vector $\vec{x}[n]$. The corresponding variations $\mathcal{V}_i(\tau_g)$ would be combined into a vector $\vec{\mathcal{V}}(\tau_g)$ to be minimized when its length is shortest. The variation components could be weighted by the importance of the respective $x_i[n]$. An analogous approach could be used with the SPEC-PEAKS algorithm in which a partial energy vector $\vec{\mathcal{E}}(\phi_g)$ would be maximized.

Rader [5] has also suggested that the vector waveform approach could be applied to the FAST-SCAN algorithm to decrease the probability of algorithm failure. All

[2]This corresponds to the incomplete time series, which in our case is the reconstructed period.

available samples would be used on each successively finer search. On earlier, coarse searches, the entire input sequence would be divided into several short sequences of length N_R, with the N_R for each iteration being determined as before, i.e., as a function of monotonically increasing Farey fraction order (Section 3.3). The same series of reduced search densities would be used as before, but more data would be utilized, thus reducing algorithm sensitivity to noise, waveform type, etc.

Also worthy of further consideration are cases in which several signals with unequal periods are superimposed, such as those in Figure 5.6. We might wish to separate and reconstruct them from a single aliased sequence. In such cases, none of the signal periods $T_{w,i}$ could be integrally related to the sampling period T_s *or* to each other. Otherwise, irreversible aliasing could occur. The Rader algorithm might be useful for determining the $T_{w,i}$ from the minima of the variation function. However, signal separation and reconstruction would be arduous (if not impossible) tasks.

Since the SPEC-PEAKS algorithm automatically eliminates all unwanted portions of the aliased spectrum, it might be more suitable for this purpose. It probably would be necessary to determine the fundamental frequency of the strongest[3] signal, reconstruct and subtract it from the aliased signal, then repeat the process. However, amplitude normalization and output sample spacing may make this approach cumbersome. These two problems also must be circumvented to permit measurement of the distortion (e.g., mean-squared error) introduced by each algorithm presented in this thesis. This should be examined, as well.

[3]The one yielding the greatest partial energy.

Bibliography

[1] C.M. Rader. Recovery of undersampled periodic waveforms. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 25(3):242–249, June 1977.

[2] R.J. Marks II. Restoration of continuously sampled band-limited signals from aliased data. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 30(5):937–942, December 1982.

[3] K. Swaminathan. Signal restoration from data aliased in time. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 33(1):151–159, February 1985.

[4] F.D. Powell. Periodic sampling of broad-band sparse spectra. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 31(5):1317–1319, October 1983.

[5] C.M. Rader. Personal communication.

[6] A.J. Jerri. The shannon sampling theorem — its various extensions and applications: a tutorial review. *Proceedings of the IEEE*, 65(11):1565–1598, November 1977.

[7] G.H. Hardy and E.M. Wright. *An Introduction to the Theory of Numbers.* Oxford University Press, London, fourth edition, 1938.

[8] O. Ore. *Number Theory and its History.* McGraw-Hill, Inc., New York, NY, 1948.

[9] N.S. Szabo and R.I. Tanaka. *Residue Arithmetic and its Applications to Computer Technology.* McGraw-Hill, Inc., New York, NY, 1967.

[10] J.H. McClellan and C.M. Rader. *Number Theory in Digital Signal Processing.* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1979.

[11] P.S. Naidu and B. Paramasivaiah. Estimation of sinusoids from incomplete time series. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 32(3):559–562, June 1984.

[12] A. Papoulis. A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Trans. Circuits and Systems*, 22(9):735–742, September 1975.

[13] R.J. Marks. Restoring lost samples from an oversampled band-limited signal. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 31(3):752–755, June 1983.

[14] N.S. Reddy and M.N.S. Swamy. Time-domain estimation of unambiguous doppler frequency in low and medium prf radars. In *ICASSP Proceedings*, pages 687–690. 1983.

[15] A.V. Oppenheim and R.W. Schafer. *Digital Signal Processing*. Prentice Hall, Englewood Cliffs, NJ, 1975.

[16] M.I. Skolnik. *Radar Handbook*. McGraw-Hill, Inc., New York, NY, 1970.

[17] P.G. Bartley. The practicality of processing undersampled waveforms. In *IEEE Southeastcon*, pages 199–200, 1981.

[18] R.T. Gregory. Residue arithmetic with rational operands. In *IEEE Symposium on Computer Arithmetic*, pages 144–145, 1981.

[19] W.K. Jenkins and C.F. Lee. Complex residue number arithmetic for digital signal processing. In *IEEE Asilomar Conference on Circuits, Systems, and Computers*, pages 480–483, 1981.

[20] S.A. Hovanessian. An algorithm for calculation of range in multiple prf radar. *IEEE Trans. Aerospace and Electronic Systems*, 12(2):287–290, March 1976.

[21] S.A. Hovanessian. Medium prf performance analysis. *IEEE Trans. Aerospace and Electronic Systems*, 18(3):286–296, May 1982.

DISTRIBUTION LIST

	DODAAD Code	
Director Defense Advanced Research Project Agency 1400 Wilson Boulevard Arlington, Virginia 22209 Attn: Program Management	HX1241	(1)
Head Mathematical Sciences Division Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217	N00014	(1)
Administrative Contracting Officer E19-628 Massachusetts Institute of Technology Cambridge, Massachusetts 02139	N66017	(1)
Director Naval Research Laboratory Attn: Code 2627 Washington, D.C. 20375	N00173	(6)
Defense Technical Information Center Bldg 5, Cameron Station Alexandria, Virginia 22314	S47031	(12)
Dr. Judith Daly DARPA / TTO 1400 Wilson Boulevard Arlington, Virginia 22209		(1)

END

DTIC

8-86